

Estándares de compra pública para sistemas de IA: la lista de verificación

El estándar sometido a su propia prueba

Mayo de 2025

Enpath · Investigación · Estándar (lista de verificación) · Pilar: Compra pública Publicado: mayo de 2025 · **Corte de la evidencia:** 25 de julio de 2026 **Idioma:** español · **Versión del estándar:** 2.0

Contenido

Parte I · Veredicto 1. Resumen ejecutivo 2. Hallazgos, reformulados 3. Por qué importa

Parte II · El estándar 4. De qué está hecho: las cuatro fuentes 5. Las cinco fases y las treinta verificaciones 6. La graduación por riesgo, y por qué funcionó

Parte III · La medición 7. El estándar medido: 27,6 % 8. El gradiente por fase: lo que se escribe y lo que se hace 9. Quién carga el costo de producir la evidencia 10. Los cuatro ceros 11. La predicción del hallazgo 3, evaluada

Parte IV · Contexto y revisión 12. La convergencia internacional, actualizada 13. Colombia: quince meses después 14. América Latina y la vía de las agencias centralizadas 15. Riesgos, revalorados

Parte V · Entrega 16. El estándar versión 2.0 17. Metodología y limitaciones 18. Predicciones falsables 19. Conclusión

Anexos A. Los treinta ítems con su medición B. Los siete cambios de la versión 2.0 C. Glosario D. Índice de figuras y tablas Referencias · Notas

Parte I · Veredicto

1. Resumen ejecutivo

Entre el diagnóstico —agosto de 2024, ocho causas de la mala compra de IA— y la solución faltaba el documento operable: el estándar que una oficina de contratación puede seguir sin ser experta en aprendizaje automático. Este informe lo entregó en mayo de 2025: **treinta verificaciones distribuidas en cinco fases** —antes del pliego, pliego, contrato, ejecución y operación—, cada una con su criterio de cumplimiento y su ámbito de aplicación según la matriz de impacto × reversibilidad.

Quince meses después, el estándar tiene algo que casi ningún instrumento de política pública tiene: **una medición de sí mismo**.

1.1 El veredicto en una tabla

Pregunta	Respuesta medida	Evidencia
¿Es medible el estándar sobre expedientes reales?	Sí	30 ítems × 20 procesos; 525 verificaciones aplicables
¿Es aplicable sin sobrecargar a la entidad?	Sí	La graduación produjo denominadores de 6 a 20; ningún ítem quedó inaplicable
¿Qué proporción se cumple?	27,6 % agregado	145 verificaciones de 525
¿Dónde se cumple y dónde no?	Se pacta y no se verifica	Fase 1: 38,0 %; fase 4: 11,7 %
¿Qué ítems concentran el poder de filtro?	Los de producción propia de la entidad	Cuatro ítems en cero, los cuatro de la entidad
¿Basta con publicar el estándar?	No	Ninguna agencia centralizada lo ha adoptado como documento tipo

Tabla 1. Lo que la medición responde. Fuente: Enpath.

1.2 El número

27,6 %. Ciento cuarenta y cinco de quinientas veinticinco verificaciones aplicables. Es el cumplimiento agregado del estándar en veinte adquisiciones reales de sistemas de IA en América Latina, medidas en noviembre de 2025 con corte documental sobre expedientes públicos.

El número exige dos lecturas simultáneas. La primera: **el mercado de compra pública de IA en la región opera, en promedio, sin tres cuartas partes de las salvaguardas que el propio consenso internacional considera mínimas.** La segunda, menos cómoda para este informe: **un estándar que la práctica cumple en 27,6 % no está midiendo una desviación; está describiendo un régimen distinto del que supone.** La distancia entre un instrumento y la realidad que pretende ordenar puede ser tan grande que el instrumento deje de ser una vara y pase a ser una aspiración. La sección 17 responde a eso sin bajar el estándar: no reduciendo la exigencia, sino ordenándola.

1.3 Los tres hallazgos que la medición produjo

Primero, el gradiente. El cumplimiento no es uniforme entre fases: pliego 38,0 %, contrato 35,8 %, antes del pliego 21,1 %, ejecución 20,2 %, operación 11,7 %. Lo que se escribe se cumple tres veces más que lo que se opera. La compra pública de IA en la región **pacta razonablemente y verifica muy poco.**

Segundo, los cuatro ceros. Cuatro verificaciones no se cumplieron en ningún proceso: la clasificación del sistema en la matriz de impacto, el contraste con la madurez de la entidad, la publicación de la ficha en el registro de algoritmos y la publicación de métricas trimestrales. No son ítems exóticos: son los cuatro que exigen que **la entidad produzca un documento propio** que antes no existía.

Tercero, la hipótesis que la medición sugiere y no demuestra. Separando los treinta ítems por quién debe producir la evidencia —diecisiete de producción propia de la entidad, trece de exigencia al proveedor— la diferencia es de 15,8 puntos porcentuales en el agregado (20,5 % frente a 36,3 %) y de 13,8 puntos en la media por ítem. Una prueba de permutación de dos colas con 200.000 remuestreos arroja **p = 0,056**: la dirección es consistente, el umbral convencional no se alcanza.

Con treinta ítems no da para llamarlo ley. Da para llamarlo hipótesis, escribirla y medirla otra vez. Es lo que hace la sección 19.

1.4 Lo que este documento entrega

El estándar **versión 2.0** conserva las treinta verificaciones y cambia siete cosas alrededor de ellas: un núcleo obligatorio de siete ítems que no admite graduación; la reasignación de los cuatro ceros a momentos del expediente donde ya existe un trámite; la sustitución del hallazgo 3 por un criterio de filtro derivado de la medición; la incorporación de las cláusulas modelo europeas en su versión de marzo de 2025; la de la Ficha de Transparencia Algorítmica chilena como formato de referencia del ítem 25; la corrección de un error de datos en la lista publicada; y una regla de verificación que faltaba: **quién comprueba que la casilla marcada corresponde a un documento que existe**.

2. Hallazgos, reformulados

Los cinco hallazgos originales se conservan cuando la evidencia los sostiene y se sustituyen cuando no. Se añaden dos que la medición produjo.

H1 · La compra es donde toda la serie se vuelve exigible — y donde casi nada se exige. El estándar es, efectivamente, el mecanismo de transmisión entre los marcos y el mercado: las siete preguntas, la ficha de datos, el protocolo de supervisión y el registro entran al expediente contractual por esta vía y por ninguna otra. Pero la medición muestra que el mecanismo transmite mal: **145 de 525 verificaciones**. Lo que la serie construyó llega al mercado en poco más de la cuarta parte de los casos.

H2 · La graduación por riesgo funcionó. Era la apuesta más frágil del diseño: aplicar los treinta ítems a un chatbot interno lo mataría por sobrecarga; aplicar diez a un sistema de fiscalización lo mataría por insuficiencia. En la medición, los denominadores efectivos por ítem fueron de **6 a 20** procesos, según el ámbito de cada verificación. Ningún ítem resultó inaplicable a todo el conjunto y ninguno se aplicó indiscriminadamente. La graduación no es un adorno del documento: es lo que hizo posible medirlo.

H3 · El poder de filtro no está donde el informe lo puso. La intuición era razonable y resultó falsa: los tres ítems designados quedaron en el montón medio. Los ítems que separan de verdad —los que ningún proceso cumplió— son otros, y no son los que un proveedor no puede cumplir: son **los que la entidad no produce**. El criterio de filtro corregido está en §11.3.

H4 · Los cuatro ceros son los cuatro documentos que nadie tenía que escribir antes. Clasificación en la matriz, contraste con la madurez, ficha en el registro y métricas trimestrales publicadas: 0 de 20, 0 de 20, 0 de 20 y 0 de 17. Los cuatro exigen que alguien dentro de la entidad redacte y firme un documento que no existía en el expediente tradicional de contratación. Ninguno depende del proveedor. Ninguno cuesta dinero. Los cuatro tienen dueño incierto.

H5 · El cumplimiento parece seguir a quién carga el costo de producir la evidencia. Los ítems que la entidad debe producir por sí misma se cumplen en 20,5 %; los que consisten en exigir algo al proveedor, en 36,3 %. La diferencia por ítem es de 13,8 puntos, con **p = 0,056** en una prueba de permutación de dos colas — por encima del umbral convencional. Los cuatro ceros caen todos en el grupo de producción propia, lo que por sí solo tiene una probabilidad de 8,7 % bajo la hipótesis

nula. Dos señales concordantes, ninguna concluyente. Se declara como hipótesis y se somete a prueba en §18.

H6 · El estándar reequilibra la negociación sin cerrar el mercado. La afirmación era que los requisitos, formulados sobre evidencia exigible y no sobre certificaciones costosas, filtran por seriedad y no por tamaño. El scorecard no midió tamaño ni origen de los proveedores, de modo que la afirmación **sigue sin prueba**. Se conserva como principio de diseño declarado y se retira de la lista de hallazgos empíricos, donde no debió estar.

H7 · Un estándar publicado no es un estándar adoptado. Quince meses después de su publicación no se localizó ninguna entidad que haya adoptado formalmente la lista como requisito de sus procesos. Colombia Compra Eficiente avanzó del silencio a un **estudio diagnóstico sobre la contratación pública de soluciones tecnológicas que incorporan inteligencia artificial**, publicado en su boletín *Radar de la compra pública*, edición N° 3 de 2026 — un paso real, y no un documento tipo. El riesgo que el informe no contempló era este: no que el estándar se cumpliera mal, sino que nadie lo intentara.

3. Por qué importa

La compra pública sigue siendo la única regulación exigible de IA en casi toda la región. Eso no cambió entre mayo de 2025 y julio de 2026; se agravó. La Unión Europea trasladó las obligaciones de alto riesgo del Anexo III al 2 de diciembre de 2027 y las del Anexo I al 2 de agosto de 2028, de modo que el efecto de arrastre que los compradores latinoamericanos esperaban del mercado europeo —proveedores que llegan con la documentación ya hecha— se retrasa dos años y medio. Perú puso en vigencia su reglamento el 22 de enero de 2026 con plazos sectoriales que empiezan a correr en septiembre. Colombia tiene el CONPES 4144 en implementación, con 106 acciones y 479.273 millones de pesos, y ninguna de esas acciones ordena un estándar de compra.

En ese vacío, **cada pliego publicado sin estándar es una vigencia contractual completa de opacidad pactada**. El scorecard permite ponerle número a la frase: en veinte adquisiciones reales, ninguna publicó la ficha del sistema, ninguna publicó métricas trimestrales, una exigió desempeño desagregado por grupos y dos exigieron ficha de datos.

Este documento apoya tres decisiones, y ahora con evidencia detrás de cada una:

- **La entidad compradora** decide qué exigir en el próximo proceso. La versión 2.0 le dice, además, **por dónde empezar**: los siete ítems del núcleo obligatorio (§16.1), elegidos porque la medición muestra que su ausencia arrastra a los demás.
- **Las agencias de compra centralizadas** deciden qué incorporar a documentos tipo e instrumentos agregados. La medición les entrega el argumento que faltaba: el cumplimiento no falla por resistencia del mercado —los ítems que dependen del proveedor se cumplen en 36,3 %— sino por **ausencia de un formato institucional para lo que la entidad debe producir**. Eso es exactamente lo que un documento tipo resuelve.
- **Los proveedores serios** deciden qué preparar de antemano. El dato que les concierne: los tres ítems más cumplidos del tablero —niveles de servicio (70 %), seguridad y continuidad (65 %), documentación técnica (60 %)— son los que ya venden. La distancia entre un proveedor serio y el estándar completo es menor de lo que ambos lados suponen.

Parte II · El estándar

4. De qué está hecho: las cuatro fuentes

El estándar no es invención sin respaldo. Cada ítem existe porque su ausencia tiene expediente. Esta edición actualiza las cuatro fuentes a su estado de julio de 2026 y corrige dos que estaban desactualizadas ya en mayo de 2025.

4.1 El diagnóstico propio

El informe de agosto de 2024 documentó ocho causas de la mala compra de IA y propuso diez cláusulas contractuales mínimas. El estándar las absorbe íntegras como fase 2, ítem 15. La medición asigna a ese ítem un cumplimiento de **6 de 20 procesos (30,0 %)** — el núcleo contractual del diagnóstico llegó a menos de un tercio del mercado examinado quince meses después de publicarse.

4.2 La convergencia internacional, en su versión vigente

Instrumento	Versión citada en mayo de 2025	Versión vigente	Efecto sobre el estándar
Cláusulas modelo de IA (UE)	29 de septiembre de 2023	5 de marzo de 2025: versión completa para alto riesgo, versión ligera, comentario; 24 lenguas	La versión ligera valida la graduación por riesgo del estándar y le da un texto contractual traducido al español para los ítems 15–20
Guía de compra de IA (Reino Unido)	Guidelines for AI procurement, 2020	AI Playbook for the UK Government , 10 de febrero de 2025 (GDS)	Amplía de la compra al ciclo completo: selección, compra, despliegue y gobierno. Refuerza las fases 3 y 4, las peor cumplidas
Práctica federal (EE. UU.)	M-24-10 y M-24-18; sucesores M-25-21 y M-25-22	M-25-21 y M-25-22 , 3 de abril de 2025	Confirmado. Categoría única de «IA de alto impacto», inventario y responsable de IA por agencia
Experiencia regional (Chile)	«Formularios de transparencia algorítmica»	Bases Tipo con requisitos éticos (enero de 2023) + Evaluación de Impacto Algorítmico + medición de sesgo y equidad + Ficha de Transparencia Algorítmica (2023–2025)	Es un cuerpo de instrumentos, no un formulario. La Ficha de Transparencia Algorítmica es el formato de referencia del ítem 25
Marco regulatorio (UE)	Reglamento (UE) 2024/1689	Modificado por el Digital Omnibus on AI : Anexo III al 2 de diciembre de 2027, Anexo I al 2 de agosto de 2028	Retrasa el arrastre de mercado dos años y medio. Aumenta, no reduce, la importancia del estándar como sustituto

Tabla 2. Las anclas internacionales del estándar, actualizadas. Fuente: Enpath sobre documentos primarios.

La entrada más relevante de esta tabla para una entidad que quiera actuar mañana es la primera: **existe, desde marzo de 2025, un texto contractual modelo en español, con versión ligera para sistemas que no son de alto riesgo y con comentario de aplicación.** Una entidad no tiene que redactar la fase 2: tiene que adaptar un texto oficial que ya existe en su idioma. Es el dato de mayor valor práctico inmediato de toda esta sección.

4.3 Los marcos de la serie 2025

El estándar es el punto donde cuatro marcos previos se vuelven exigibles: el modelo de madurez de febrero entra como filtro de fase 0 (ítem 3); la ficha de datos de marzo, como requisito de fase 1 (ítem 9); el registro de abril, como obligación de fase 3 (ítem 25); y el protocolo de supervisión del marco de enero, como condición de despliegue (ítem 23). Los cuatro se miden. Los cuatro resultados aparecen en la tabla 3.

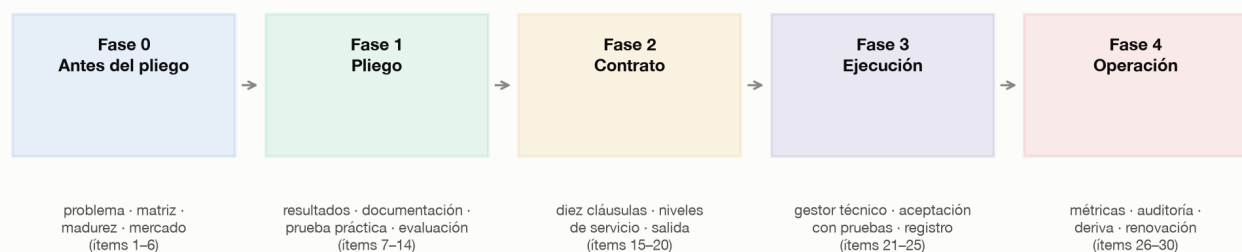
Marco de la serie	Ítem que lo incorpora	Cumplimiento medido	Lectura
Modelo de madurez (feb. 2025)	3 · Madurez contrastada, regla de ambición	0 de 20 (0 %)	Nadie contrastó su ambición con su capacidad
Marco de datos (mar. 2025)	9 · Ficha de datos exigida	2 de 14 (14,3 %)	El componente 7 del marco llegó a dos procesos
Registro de algoritmos (abr. 2025)	25 · Ficha publicada en el registro	0 de 20 (0 %)	El registro no existe como práctica contractual
Marco de gobernanza (ene. 2025), puerta 3	23 · Protocolo de supervisión aprobado	3 de 14 (21,4 %)	La puerta 3 se cruzó en uno de cada cinco

Tabla 2. Los marcos de la serie, medidos a través del estándar. Fuente: Enpath, scorecard de noviembre de 2025.

Es la tabla más severa de este documento y no admite reinterpretación favorable: **la serie 2025 de Enpath, medida por su propio instrumento en el terreno donde debía volverse exigible, promedia menos de 9 % de cumplimiento**. Publicar un marco y publicar el instrumento que lo hace exigible no basta. El informe de abril de 2025 sobre registros llegó a la misma conclusión por otra vía y con la misma honestidad; el de marzo, también. Tres mediciones independientes convergen en el mismo diagnóstico, y ese diagnóstico es sobre el emisor, no sobre el receptor.

5. Las cinco fases y las treinta verificaciones

Estándar Enpath de compra de IA · Treinta verificaciones en cinco fases



Fuente: Enpath. Lista completa en [data/lista_de_verificacion.csv](#); cada ítem indica si aplica a todos los sistemas o solo a impacto medio/alto.

Figura 1. Las cinco fases y sus treinta verificaciones. Lista completa en [data/lista_de_verificacion.csv](#). Fuente: Enpath.

El contenido de las cinco fases se conserva sin cambios respecto de mayo de 2025. Lo que sigue añade, a cada una, el número que la medición produjo.

5.1 Fase 0 · Antes del pliego (ítems 1–6) — 24 de 114 · 21,1 %

La fase que decide todo y no aparece en ningún expediente tradicional: problema por escrito, clasificación en la matriz, contraste con la madurez, alternativas no algorítmicas evaluadas, diálogo de mercado sobre el problema —nunca sobre una solución— y base legal de datos verificada. **Regla de oro conservada:** si la fase 0 no está documentada, el proceso no se abre.

Su medición contiene el resultado más revelador del tablero: dentro de la misma fase conviven el ítem que mejor se cumple de los seis (base legal verificada, 9 de 20) y dos de los cuatro ceros. La diferencia entre ellos no es de dificultad ni de costo — es de **existencia previa del trámite**. La verificación de base legal ya tenía dueño en la entidad: la oficina jurídica. La clasificación en la matriz y el contraste de madurez no tenían dueño en ninguna parte.

5.2 Fase 1 · Pliego (ítems 7–14) — 52 de 137 · 38,0 %

El pliego compra resultados: objeto por métricas, documentación técnica y ficha de datos como requisitos habilitantes, desempeño por grupos cuando el sistema afecta personas, prueba práctica sobre datos de la entidad como criterio ponderado, comité con perfil técnico, seguridad y continuidad de terceros, y costo total de la vida útil en la oferta.

Es la fase mejor cumplida, y con distancia interna enorme: seguridad y continuidad, 13 de 20; documentación técnica, 12 de 20; pero costo total de vida útil, 3 de 20, y desempeño por grupos, 1 de 9. Lo que el pliego ya sabía pedir se pide; lo que el estándar añadió, casi no.

La exigencia sigue al riesgo · Requisitos del pliego según la matriz de impacto

	Bajo impacto	Impacto medio	Alto impacto
Documentación técnica	Resumida	Completa	Completa
Ficha de datos	Simplificada	Completa	Completa y verificada
Prueba práctica en evaluación	Opcional	Obligatoria	Obligatoria y ampliada
Desempeño por grupos	—	Si afecta personas	Obligatorio
Auditoría externa	—	Según contrato	Anual obligatoria
Protocolo de supervisión	Monitoreo (patrón C)	Cola de excepciones (B)	Revisión previa (A)

Fuente: Enpath. La clasificación proviene de la matriz de delegación (febrero de 2024).

Figura 2. La graduación de requisitos según la matriz de impacto. Fuente: Enpath.

5.3 Fase 2 · Contrato (ítems 15–20) — 43 de 120 · 35,8 %

Las diez cláusulas de agosto de 2024, más niveles de servicio sobre métricas, hitos de validación verificables, exportación con formato pactado, gestión de cambios de modelo y plan de salida con transición asistida.

Aquí aparece el ítem más cumplido de los treinta —niveles de servicio, 14 de 20 (70 %)— junto a uno de los menos cumplidos —gestión de cambios de modelo, 2 de 20 (10 %)—. La lectura es directa: **el contrato de servicios de tecnología ya existía y el contrato de sistemas que aprenden todavía no.**

Un nivel de servicio sobre disponibilidad es una cláusula que cualquier oficina jurídica sabe redactar; una cláusula que somete el cambio de modelo a aprobación previa de la entidad exige entender que el objeto contratado cambia solo.

5.4 Fase 3 · Ejecución (ítems 21–25) — 19 de 94 · 20,2 %

Gestor de contrato con capacidad técnica, aceptación contra pruebas pactadas —nunca contra demostración—, protocolo de supervisión aprobado antes de producción, línea de base medida y ficha publicada en el registro.

La aceptación es el momento de mayor riesgo de captura: lo que se acepta mal se opera mal durante toda la vigencia. Se cumplió en 6 de 20 procesos. La línea de base —medir cómo funcionaba el proceso *antes* del sistema, sin lo cual ninguna mejora es demostrable— se midió en 2 de 20. Y la ficha del registro, en ninguno.

5.5 Fase 4 · Operación (ítems 26–30) — 7 de 60 · 11,7 %

Métricas trimestrales publicadas, auditoría según calendario, monitoreo de deriva con umbrales, renovación únicamente por puerta 4 y cierre con exportación verificada. La fase que convierte el contrato en gobernanza continua — y cuya omisión explica los pilotos eternos y las renovaciones por inercia.

Siete verificaciones cumplidas de sesenta aplicables. Es el número que mejor resume el estado de la compra pública de IA en la región: **se compra, y después no se mira.**

6. La graduación por riesgo, y por qué funcionó

La graduación era la apuesta de diseño más frágil del estándar. Su lógica: la exigencia sigue a la matriz de impacto × reversibilidad, no al valor del contrato — un sistema barato de alto impacto exige más que uno caro de bajo impacto.

La medición permite evaluarla porque el scorecard registró, para cada ítem, cuántos de los veinte procesos le eran **aplicables**.

Universo aplicable	Ítems	Ejemplos
20 de 20 (todos)	16	Problema por escrito, clasificación en matriz, niveles de servicio, ficha en registro
17 de 20	1	Métricas trimestrales publicadas
14 de 20	5	Ficha de datos, prueba práctica, protocolo de supervisión, monitoreo de deriva, diálogo de mercado
12 de 20	1	Auditoría externa
11 de 20	1	Renovación con evaluación documentada
9 de 20	1	Desempeño desagregado por grupos
6 de 20	1	Cierre con exportación verificada

Tabla 3. La graduación en la práctica: cuántos procesos quedaron sujetos a cada ítem. Fuente: Enpath, scorecard de noviembre de 2025.

Tres conclusiones.

Primera: la graduación operó. De 600 pares ítem-proceso posibles, 525 resultaron aplicables — un 87,5 %. La reducción del 12,5 % no es trivial ni es ruido: recae sobre los ítems más costosos, que es exactamente donde debía recaer.

Segunda: la graduación no fue una excusa. El temor legítimo con cualquier esquema de exigencia graduada es que las entidades se autoclasifiquen a la baja para escapar de los requisitos duros. Aquí no pudo ocurrir, porque **el evaluador fue externo y clasificó sobre el expediente**. Es una advertencia para la adopción real: en autoevaluación, el ítem 2 —clasificación en la matriz— pasa de ser un requisito a ser la puerta por la que se sale de todos los demás. La versión 2.0 responde con la regla de verificación de §16.6.

Tercera: el cumplimiento no se explica por el tamaño del universo. Los ítems de universo reducido no se cumplen mejor que los universales: entre los de 14 procesos aplicables hay uno al 7,1 % y otro al 35,7 %. La graduación repartió carga; no repartió cumplimiento.

Parte III · La medición

7. El estándar medido: 27,6 %

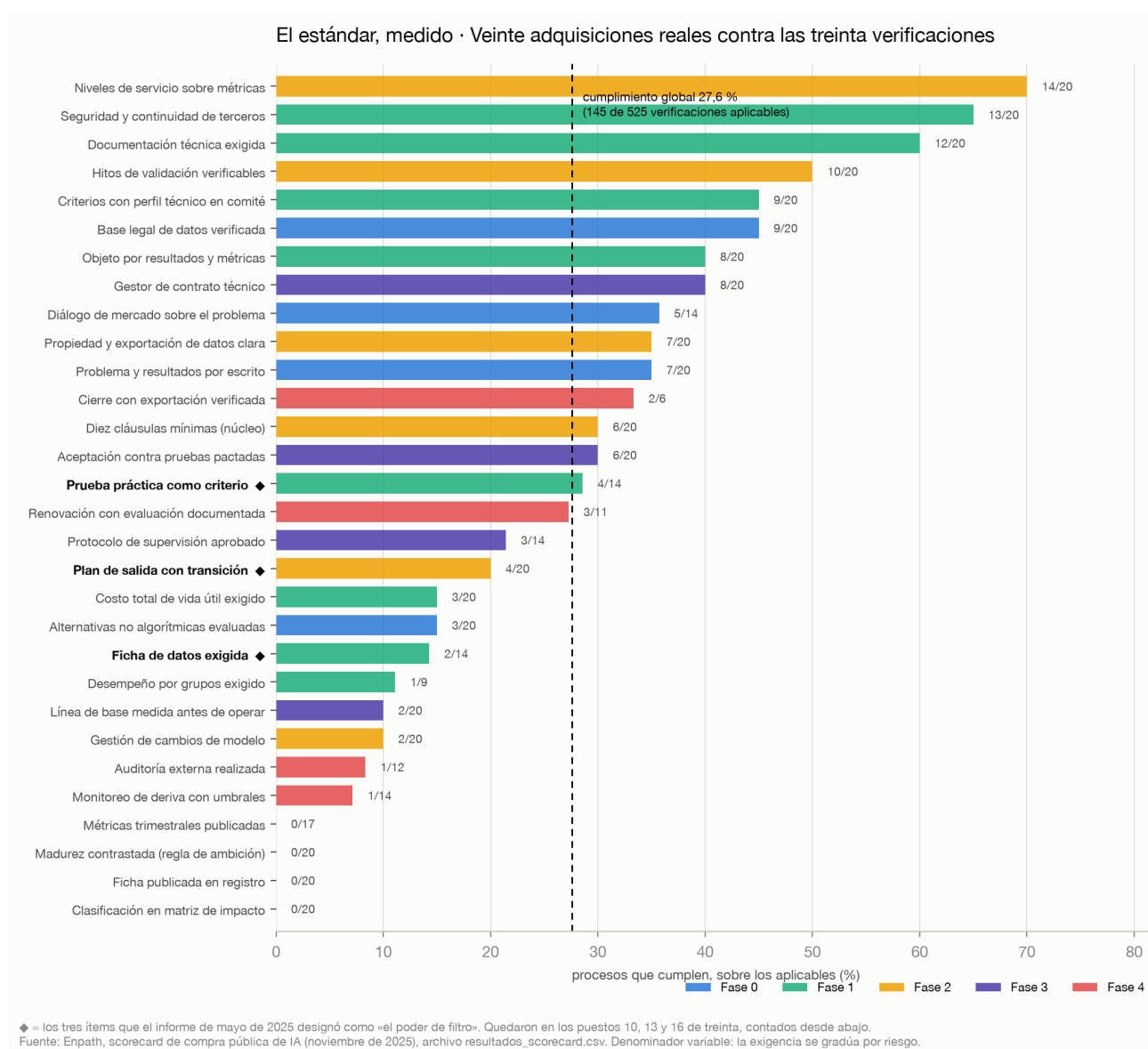


Figura 3. Los treinta ítems del estándar ordenados por cumplimiento, sobre veinte adquisiciones reales de América Latina. Los rombos marcan los tres ítems que la intuición señala como «el poder de filtro». Fuente: Enpath, scorecard de noviembre de 2025.

145 de 525 verificaciones aplicables: 27,6 %.

El tablero completo, con sus treinta filas, permite ver tres cosas que el número agregado esconde.

Arriba, lo que el mercado ya vendía. Los cuatro ítems mejor cumplidos —niveles de servicio (70 %), seguridad y continuidad de terceros (65 %), documentación técnica (60 %), hitos de validación (50 %)— son verificaciones que existían en la contratación de tecnología antes de que existiera la IA. Ninguna es una novedad del estándar. El estándar no las creó: las heredó.

Abajo, lo que el estándar añadió. Los diez ítems peor cumplidos incluyen los cuatro ceros, el monitoreo de deriva (7,1 %), la auditoría externa (8,3 %), la gestión de cambios de modelo (10 %), la

línea de base (10 %), el desempeño por grupos (11,1 %) y la ficha de datos (14,3 %). Son, uno por uno, los requisitos específicos de un objeto que aprende, cambia y afecta personas.

En medio, casi todo lo demás. Entre el 15 % y el 45 % se acumulan diecisiete de los treinta ítems, sin agrupamiento claro por fase. Es la zona donde el cumplimiento depende de la entidad concreta y no del ítem — y por eso es la zona donde un documento tipo tendría el mayor efecto marginal.

7.1 Lo que el número no dice

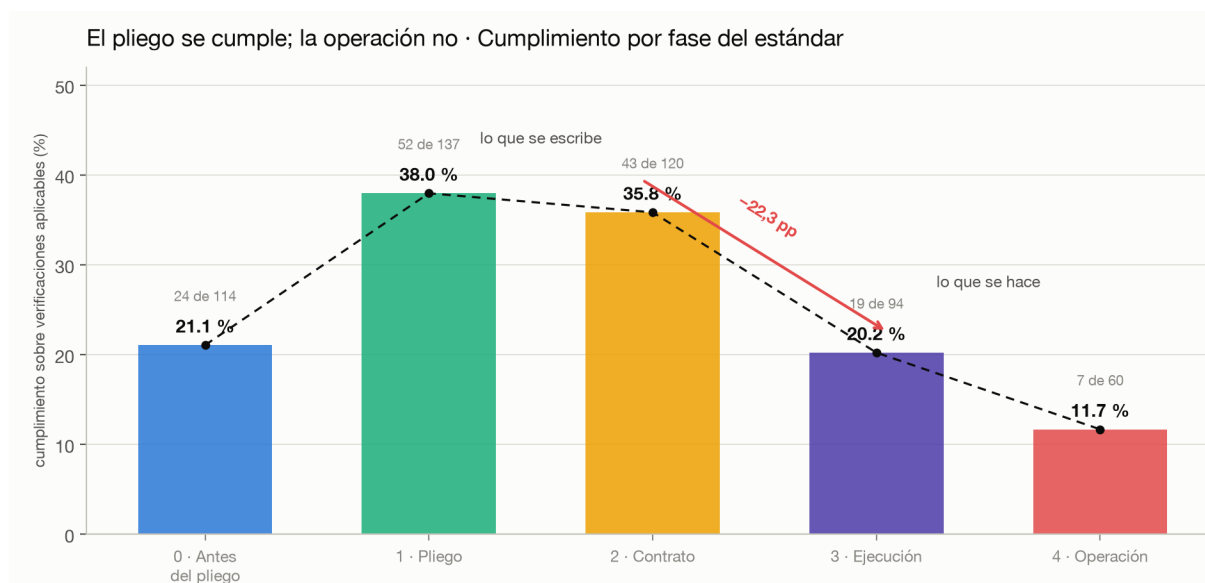
Tres advertencias, todas del lado de la prudencia.

El scorecard midió **veinte adquisiciones seleccionadas**, no una muestra aleatoria del universo de compras públicas de IA en la región. El criterio de selección está declarado en su metodología. Si la selección privilegió procesos visibles y documentados —lo habitual cuando se depende de expedientes públicos—, **27,6 % es un techo, no un promedio.**

La medición es **documental**: verifica lo que el expediente muestra. Un proceso que hizo bien una cosa y no la documentó cuenta como incumplimiento. Es la elección correcta para un estándar cuya premisa es que lo no documentado no es exigible, pero conviene decirlo: el número mide la transparencia del cumplimiento tanto como el cumplimiento.

Y la medición es **de un momento**. Los expedientes examinados corresponden a procesos anteriores a noviembre de 2025, muchos de ellos anteriores a la publicación del estándar. Ningún proceso del scorecard pudo haber usado el estándar como guía. Lo que se midió no es «cuánto cumple quien lo intenta», sino **cuánto coincide la práctica espontánea con el consenso internacional**. La respuesta —poco más de un cuarto— sigue siendo el dato relevante para quien decide si publica un documento tipo.

8. El gradiente por fase: lo que se escribe y lo que se hace



Fuente: Enpath, scorecard de noviembre de 2025. La caída entre el pliego (38,0 %) y la ejecución (20,2 %) es la brecha entre pactar y verificar; la fase 4 (11,7 %) es la brecha entre operar y rendir cuentas.

Figura 4. Cumplimiento por fase del estándar. Fuente: Enpath, scorecard de noviembre de 2025.

Fase	Ítems	Verificaciones cumplidas	Aplicables	Cumplimiento
0 · Antes del pliego	6	24	114	21,1 %
1 · Pliego	8	52	137	38,0 %
2 · Contrato	6	43	120	35,8 %
3 · Ejecución	5	19	94	20,2 %
4 · Operación	5	7	60	11,7 %
Total	30	145	525	27,6 %

Tabla 4. Cumplimiento por fase. Fuente: Enpath, scorecard de noviembre de 2025.

La forma del gradiente es lo importante: **sube hasta el pliego y cae después**. Entre la fase 1 y la fase 4 hay 26,3 puntos porcentuales de caída; entre el pliego y la ejecución, 17,8 puntos.

Esto refuta una parte del hallazgo 1 original. El informe sostuvo que «la compra es donde toda la serie se vuelve exigible», con la implicación de que asegurar el pliego aseguraba el resto. La medición dice otra cosa: **el pliego es donde la exigencia se escribe, y la exigencia escrita no se transmite sola a la ejecución**. Un proceso que exigió prueba práctica en el pliego puede aceptar contra demostración en la ejecución; un contrato con niveles de servicio sobre métricas puede operar sin que nadie publique esas métricas. Hay 52 verificaciones cumplidas en el pliego y 7 en la operación.

8.1 La brecha pactar-verificar

La caída tiene una explicación estructural y no motivacional. Las fases 0, 1 y 2 se ejecutan **dentro del proceso contractual**, con expediente, plazos, responsable identificado y consecuencias jurídicas del incumplimiento. Las fases 3 y 4 se ejecutan **después de la firma**, cuando el proceso contractual ya terminó y la supervisión pasa a un supervisor de contrato cuya carga formal es verificar entregas y autorizar pagos, no evaluar el desempeño de un modelo.

En términos institucionales: **la fase 1 tiene dueño, la fase 4 no**. El ítem 21 —gestor de contrato con capacidad técnica— es el que debía cerrar esa brecha, y se cumple en 8 de 20. Es el punto de apalancamiento del estándar y la versión 2.0 lo asciende al núcleo obligatorio.

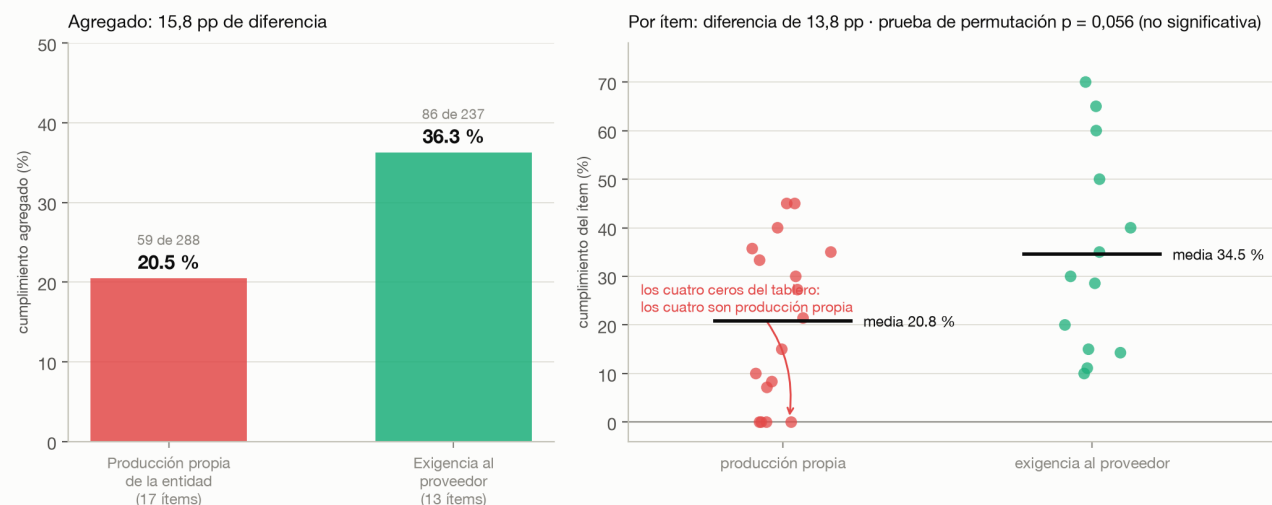
8.2 La fase 0, o el precio de no tener trámite

La fase 0 rompe el patrón: es previa al proceso, y sin embargo se cumple en 21,1 %, apenas por encima de la ejecución. La razón es la misma vista desde el otro lado: **la fase 0 tampoco tiene expediente**. Ocurre antes de que exista número de proceso, antes de que exista carpeta, antes de que exista responsable formal. Es una fase de trabajo institucional sin recipiente administrativo donde depositarse.

Sus dos ítems al 0 % son precisamente los que exigen crear un documento nuevo; sus dos ítems mejor cumplidos —base legal verificada (45 %) y diálogo de mercado (35,7 %)— son los que se apoyan en trámites que ya existían: el concepto jurídico y el estudio de mercado. El diseño de la versión 2.0 sale directamente de esta observación: **no pedir documentos nuevos, sino ampliar los que ya se producen**.

9. Quién carga el costo de producir la evidencia

El cumplimiento sigue a quién paga el costo de producir la evidencia



Fuente: Enpath, scorecard de noviembre de 2025 (n = 30 ítems). Prueba de permutación de dos colas sobre la diferencia de medias por ítem, 200.000 resamples, semilla 20250531. La dirección es consistente en las dos formulaciones, pero con treinta ítems no alcanza el umbral convencional: es una hipótesis para la próxima medición, no una ley demostrada.

Figura 5. Cumplimiento según quién debe producir la evidencia. Fuente: Enpath, scorecard de noviembre de 2025.

Una clasificación adicional ordena el resultado. Cada uno de los treinta ítems se asigna a uno de dos grupos según **quién debe producir el documento o la evidencia** que lo satisface:

- **Producción propia de la entidad (17 ítems).** La entidad debe generar un análisis, una decisión documentada o una publicación que antes no existía: la clasificación en la matriz, el contraste de madurez, las alternativas evaluadas, el diálogo de mercado, la verificación de base legal, el problema por escrito, el comité con perfil técnico, el gestor técnico, la aceptación contra pruebas, el protocolo de supervisión, la línea de base, la ficha en el registro, las métricas trimestrales publicadas, la auditoría externa, la renovación evaluada, el cierre con exportación verificada y el monitoreo de deriva.
- **Exigencia al proveedor (13 ítems).** La entidad escribe un requisito y el proveedor lo satisface: documentación técnica, ficha de datos, desempeño por grupos, prueba práctica, seguridad y continuidad, costo de vida útil, objeto por resultados, las diez cláusulas, niveles de servicio, hitos de validación, propiedad y exportación, gestión de cambios de modelo y plan de salida.

Grupo	Ítems	Cumplidas	Aplicables	Cumplimiento agregado	Media por ítem
Producción propia de la entidad	17	59	288	20,5 %	20,8 %
Exigencia al proveedor	13	86	237	36,3 %	34,5 %
Diferencia	—	—	—	15,8 pp	13,8 pp

Tabla 5. Cumplimiento según quién produce la evidencia. Fuente: Enpath, scorecard de noviembre de 2025.

9.1 La prueba, y su resultado honesto

Una diferencia de quince puntos entre dos grupos de tamaño desigual puede ser un artefacto de qué ítems cayeron en cada grupo. La prueba: una **permutación de dos colas** sobre la diferencia de

medias por ítem, reasignando al azar las etiquetas de grupo 200.000 veces (semilla 20250531) y contando cuántas veces la diferencia absoluta simulada iguala o supera la observada.

$p = 0,056$.

Por encima del umbral convencional de 0,05. La diferencia observada no es despreciable, pero con treinta ítems la evidencia no permite descartar el azar con la confianza que la afirmación merecería. Reportarlo como significativo sería exactamente el tipo de sobreinterpretación que este informe le exige a los proveedores no hacer con sus métricas de desempeño.

Una segunda señal, independiente de la primera: **los cuatro ceros del tablero caen los cuatro en el grupo de producción propia**. Bajo la hipótesis nula de asignación aleatoria, la probabilidad de que las cuatro peores posiciones correspondan a un grupo de 17 de 30 es de 8,7 %. También por encima del umbral, también en la misma dirección.

Dos señales concordantes que no cruzan la línea. La conclusión correcta no es «el cumplimiento sigue a quién carga el costo»; es: **la hipótesis de que el cumplimiento sigue a quién carga el costo sobrevivió a su primera prueba sin ser confirmada, y merece una segunda con más datos**. La sección 19 la escribe como predicción falsable con umbral y fecha.

9.2 Por qué la hipótesis es plausible aunque no esté probada

El mecanismo que la haría cierta es mundano. Exigirle algo al proveedor tiene, para la entidad, un costo marginal cercano a cero: es un párrafo más en un pliego que de todos modos hay que escribir, y el trabajo lo hace otro. Producir la evidencia uno mismo tiene un costo real: horas de un funcionario que no las tiene, un criterio técnico que la entidad quizá no posee, y —lo decisivo— **una firma**. Clasificar un sistema como de alto impacto es un acto que compromete a quien lo firma; no clasificarlo no compromete a nadie.

Esa asimetría no es cinismo administrativo: es la respuesta racional de una institución cuyo sistema de incentivos castiga el error visible y no castiga la omisión. El estándar de mayo de 2025 no la contempló. La versión 2.0 la contempla de la única forma que funciona: **eliminando el documento nuevo y adosando el requisito a un trámite que ya tiene dueño y firma** (§16.2).

10. Los cuatro ceros

Ítem	Fase	Resultado	Qué exige	Quién lo firmaría hoy
2 · Clasificación en la matriz de impacto	0	0 de 20	Una página que sitúe el sistema en impacto × reversibilidad	Nadie designado
3 · Madurez contrastada (regla de ambición)	0	0 de 20	Un contraste entre lo que se compra y lo que la entidad puede operar	Nadie designado
25 · Ficha publicada en el registro	3	0 de 20	Siete campos publicados al entrar en operación	Nadie designado
26 · Métricas trimestrales publicadas	4	0 de 17	Publicar cada trimestre lo que el contrato ya obliga a medir	Nadie designado

Tabla 6. Los cuatro ceros. Fuente: Enpath, scorecard de noviembre de 2025.

Los cuatro comparten cuatro rasgos y ninguno es el que se esperaría.

No son caros. Ninguno requiere contratar a nadie ni comprar nada. Los cuatro juntos son, generosamente estimados, entre dos y tres días de trabajo de un profesional por sistema.

No son técnicos. Ninguno exige entender aprendizaje automático. Clasificar un sistema en una matriz de dos ejes, contrastar ambición con capacidad, llenar siete campos y publicar cuatro números al trimestre son tareas de gestión, no de ciencia de datos.

No dependen del proveedor. Los cuatro son actos unilaterales de la entidad. Ningún proveedor puede impedirlos y ninguno se beneficia de que no ocurran.

No tienen dueño. Los cuatro exigen que alguien decida que le corresponden. En un expediente de contratación estándar no hay casilla, no hay plazo, no hay responsable y no hay consecuencia por omitirlos. Existen solo en el documento que los propuso.

El cuarto rasgo es la explicación. Y explica también por qué el ítem 6 —verificar la base legal del tratamiento de datos, que es más difícil, más técnico y más riesgoso que los cuatro anteriores— se cumple en 9 de 20: porque **la oficina jurídica ya existe, ya firma conceptos y ya tiene el trámite**. La diferencia entre 45 % y 0 % no es de dificultad. Es de domicilio institucional.

11. La predicción del hallazgo 3, evaluada

11.1 Lo que se predijo

«La experiencia comparada y el catálogo de fallas coinciden: la prueba práctica (ítem 11), la ficha de datos (ítem 9) y el plan de salida (ítem 20) son los tres requisitos que los malos proveedores no pueden cumplir y los buenos cumplen sin drama. Una entidad que solo pudiera exigir tres, exigiría esos.» — Informe original, hallazgo 3, mayo de 2025

La afirmación tiene dos partes separables. Una **normativa**: si tuvieras que elegir tres, elige estos. Y una **empírica**, implícita pero necesaria para sostener la primera: estos tres discriminan, es decir, separan a los procesos serios de los demás mejor que los otros veintisiete.

11.2 Lo que se midió

Ítem designado	Cumplimiento	Puesto de 30, desde abajo
9 · Ficha de datos exigida	2 de 14 · 14,3 %	10
20 · Plan de salida con transición	4 de 20 · 20,0 %	13
11 · Prueba práctica como criterio	4 de 14 · 28,6 %	16

Tabla 7. Los tres ítems de filtro, medidos. Fuente: Enpath, scorecard de noviembre de 2025.

Los tres quedaron en el tercio medio del tablero. Nueve ítems se cumplieron menos que la ficha de datos; catorce, menos que la prueba práctica. **La parte empírica de la afirmación es falsa**. Los tres

ítems designados no son los que más separan; son ítems de dificultad intermedia, cumplidos por entre uno y tres de cada diez procesos.

Y la nota de actualización de este informe, escrita en julio de 2026, afirmó lo contrario de lo que su propia fuente decía: que el scorecard había confirmado a estos tres como «los tres ítems menos cumplidos». No lo hizo. Esa nota se retira. Es un caso ejemplar del error que el informe de octubre de 2024 llamó *coherencia retroactiva*: leer un dato nuevo como confirmación de lo que ya se creía.

11.3 El criterio de filtro, reformulado sobre la evidencia

La pregunta del hallazgo 3 sigue siendo la correcta —¿qué exigir si solo puedes exigir unas pocas cosas?— y merece una respuesta derivada de la medición y no de la intuición. La medición ofrece dos criterios distintos, y conviene no confundirlos:

Criterio A · Los ítems cuya ausencia arrastra a los demás. Un ítem que, si falta, hace imposible o inverificable el cumplimiento de otros. La clasificación en la matriz (ítem 2) es el caso puro: sin ella, los diez ítems de exigencia graduada no tienen umbral, y la entidad no puede saber si le corresponden. La línea de base (ítem 24) es el segundo: sin ella, ni las métricas de la fase 4 ni la evaluación de renovación tienen contra qué compararse. El gestor técnico (ítem 21) es el tercero: sin él, las cinco verificaciones de las fases 3 y 4 carecen de ejecutor.

Criterio B · Los ítems que el proveedor no puede fingir. Es el criterio intuitivo, y su mejor candidato no es el que suele suponerse. No es la prueba práctica —que puede diseñarse blanda— sino la **aceptación contra pruebas pactadas** (ítem 22): la prueba práctica se pacta antes y puede negociarse; la aceptación ocurre después, con el sistema construido, y o pasa o no pasa. Se cumple en 6 de 20.

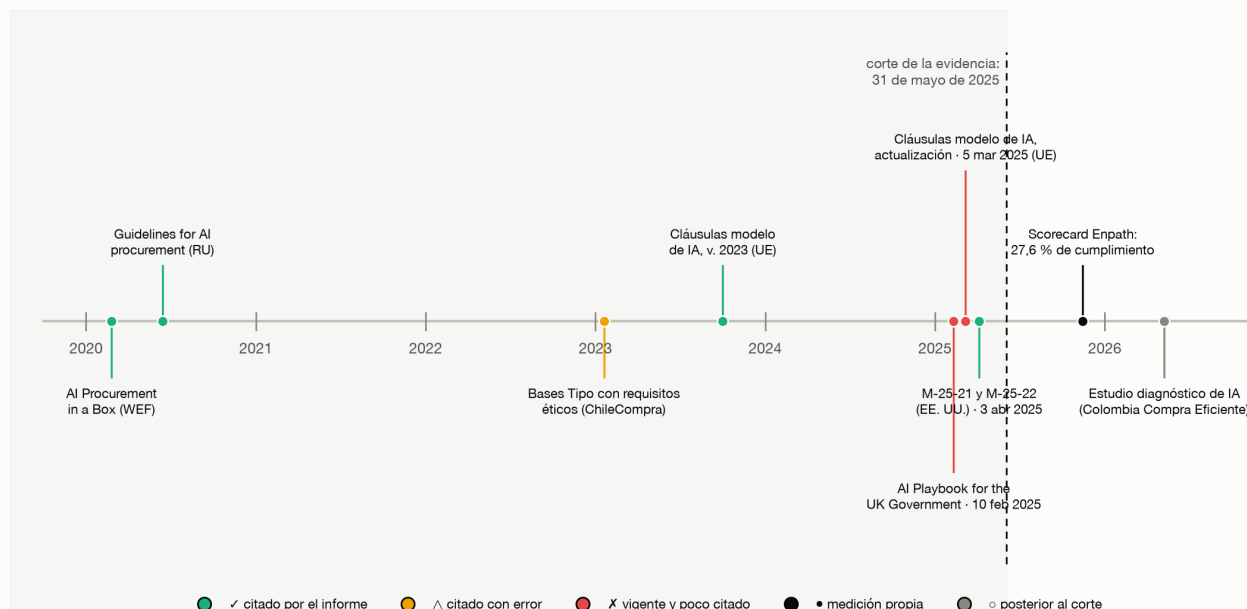
La respuesta corregida a la pregunta original. Si una entidad solo pudiera exigir tres cosas, la evidencia recomienda: el ítem 2 (clasificar), el ítem 21 (designar un gestor con capacidad técnica) y el ítem 22 (aceptar contra pruebas, nunca contra demostración). Uno abre el estándar, otro le da ejecutor y el tercero le da un momento de verdad. Los tres son de producción propia — que es, precisamente, lo que los hace difíciles y por lo que hay que nombrarlos primero.

Conviene subrayar en qué consiste el desplazamiento: **la intuición busca el filtro en lo que el proveedor debe traer, y la evidencia lo sitúa en lo que la entidad debe poner.**

Parte IV · Contexto y revisión

12. La convergencia internacional, actualizada

La convergencia internacional que el informe describió, y la que ya existía



Fuente: Enpath sobre documentos primarios. Las versiones vigentes de las cuatro anclas internacionales del estándar, con su fecha de publicación.

Figura 6. Los instrumentos internacionales de compra de IA y su fecha de vigencia. Fuente: Enpath sobre documentos primarios.

La figura muestra lo que la sección 2 registra en prosa: dos de las cuatro anclas internacionales del informe habían sido actualizadas o reemplazadas antes de su corte del 31 de mayo de 2025, y el informe citó las versiones antiguas.

12.1 Las cláusulas modelo europeas, versión de marzo de 2025

El 5 de marzo de 2025 la Comisión Europea publicó la actualización de las **cláusulas modelo de contratación de IA (MCC-AI)**, alineadas con el Reglamento (UE) 2024/1689 adoptado el 13 de junio de 2024. Sustituyen a la versión del 29 de septiembre de 2023 y se componen de tres documentos: una **versión completa** para sistemas de alto riesgo que sigue las obligaciones del capítulo III del Reglamento, una **versión ligera** personalizable para sistemas que no son de alto riesgo, y un **comentario** con orientación de uso, personalización y aplicación práctica. Los tres están disponibles en veinticuatro lenguas, español incluido.¹

Para el estándar Enpath esto tiene tres consecuencias operativas, todas favorables:

Primera: la fase 2 ya tiene texto. Los ítems 15 a 20 —las diez cláusulas mínimas, niveles de servicio, hitos, propiedad y exportación, gestión de cambios y plan de salida— piden a la entidad que redacte. Desde marzo de 2025 existe un modelo oficial en español que cubre la mayor parte de ese contenido. La versión 2.0 del estándar remite a él explícitamente.

Segunda: la graduación queda validada por un tercero. Que la Comisión publique una versión completa y una ligera según el sistema sea o no de alto riesgo es exactamente la arquitectura del

estándar Enpath: la exigencia sigue al riesgo. Que dos ejercicios independientes lleguen a la misma solución es evidencia de que la solución es la correcta.

Tercera: el argumento con el proveedor cambia de naturaleza. Pedirle a un proveedor internacional que acepte cláusulas redactadas por una entidad latinoamericana es una negociación. Pedirle que acepte las cláusulas modelo de la Comisión Europea, traducidas oficialmente y ya usadas por compradores públicos europeos, es otra cosa. La asimetría de negociación que el informe de agosto de 2024 documentó se reduce cuando el texto no es de la entidad sino del mercado más exigente.

12.2 El AI Playbook británico

El *Artificial Intelligence Playbook for the UK Government* fue publicado por el Government Digital Service el 10 de febrero de 2025. Reemplaza y amplía el marco de IA generativa de enero de 2024 y ordena en diez principios la selección, compra, despliegue y gobierno de sistemas de IA por el gobierno central y los organismos en su alcance, con controles prácticos y orientación de contratación y aseguramiento.²

Su relevancia para este estándar no es la de un competidor sino la de un complemento en la zona exacta donde el estándar se cumple peor. Las *Guidelines for AI procurement* de 2020, que el informe citó, terminaban donde termina la compra. El Playbook continúa: despliegue, operación y gobierno. Es decir, cubre las fases 3 y 4 —20,2 % y 11,7 %— con material que en mayo de 2025 ya estaba disponible.

12.3 El mapa, en una tabla

	Estándar Enpath (2025)	MCC-AI (UE, mar. 2025)	AI Playbook (RU, feb. 2025)	M-25-21/22 (EE. UU., abr. 2025)	Instrumentos chilenos (2023–2025)
Fase 0 · antes del pliego	6 ítems	—	Principios y evaluación previa	Categoría de alto impacto	Evaluación de Impacto Algorítmico
Fase 1 · pliego	8 ítems	Requisitos derivados	Orientación de contratación	Requisitos mínimos de adquisición	Bases Tipo con requisitos éticos
Fase 2 · contrato	6 ítems	Texto contractual completo, 2 versiones, 24 lenguas	—	—	Bases Tipo
Fase 3 · ejecución	5 ítems	Obligaciones de despliegue	Controles de despliegue	Responsable de IA de la agencia	Medición de sesgo y equidad
Fase 4 · operación	5 ítems	Monitoreo poscomercialización	Gobierno continuo	Inventario obligatorio	Ficha de Transparencia Algorítmica
Publicación de la ficha del sistema	Obligatoria	Registro europeo para alto riesgo	Vía ATRS	Inventario público	Ficha publicada

Tabla 8. El estándar Enpath frente a los cuatro cuerpos de referencia, en su versión vigente. Fuente: Enpath sobre documentos primarios.

Hay una diferencia deliberada, y tiene un matiz: la publicación de la ficha del sistema (ítem 25) es obligatoria en el estándar Enpath para **todos** los sistemas, mientras que el régimen europeo la vincula al alto riesgo y el estadounidense al inventario de agencia. Chile, por su parte, publicó la Fi-

cha de Transparencia Algorítmica como bien público entre 2023 y 2025: es el formato regional más cercano a lo que el ítem 25 exige y se adopta aquí como referencia.

12.4 El retraso europeo y lo que implica

El *Digital Omnibus on AI* —adoptado 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 2028.³

La expectativa razonable era que la oferta global viniera ajustada al reglamento europeo, de modo que pedir la documentación del estándar fuese pedir lo que los proveedores serios ya producen para otros mercados. Ese arrastre se retrasa. **La consecuencia es contraintuitiva y conviene decir-la claro: el retraso europeo no reduce la urgencia del estándar regional; la aumenta.** Durante dos años y medio adicionales, lo que un comprador latinoamericano no exija por contrato no se lo va a exigir nadie más.

13. Colombia: quince meses después

El informe propuso tres vehículos para convertir el estándar en regla colombiana sin reforma legal. Este es su estado.

Vehículo propuesto	Estado en julio de 2026	Evidencia
Documentos tipo de Colombia Compra Eficiente para adquisición de sistemas de decisión automatizada	No adoptado. La agencia publicó un estudio diagnóstico sobre la contratación pública de soluciones tecnológicas que incorporan inteligencia artificial en su boletín Radar de la compra pública, edición N° 3 de 2026	Boletín institucional ⁴
Condiciones de entrada en instrumentos agregados	No localizado	Búsqueda documental
Adopción unilateral por entidad mediante resolución	No localizada en ninguna entidad	Búsqueda documental; IGIA 2026
Marca de contratos de IA en SECOP	No localizada	Búsqueda documental

Tabla 9. Los tres vehículos colombianos, quince meses después. Fuente: Enpath.

El estudio diagnóstico de Colombia Compra Eficiente merece una lectura precisa, sin celebrarlo de más ni descartarlo. **Es un avance real:** significa que la agencia central identificó la compra de IA como una categoría que requiere análisis propio, lo cual es la condición previa de cualquier instrumento. **Y no es lo que se recomendó:** un diagnóstico describe el problema; un documento tipo lo resuelve para todas las entidades a la vez. La distancia entre ambos es la distancia entre saber y ordenar, y en compra pública esa distancia se mide en años.

El contexto lo hace más notable. El Índice Enpath de gobernanza de IA del sector público, publicado en marzo de 2026 sobre dieciocho ministerios, encontró acto de gobernanza en cuatro, inventario publicado en dos, ficha pública de sistema en dos, y cero en métricas publicadas, evaluación de riesgo publicada y vía de impugnación específica.⁵ Los ceros del IGIA y los ceros del scorecard son los mismos ceros vistos desde dos ángulos: el de la entidad y el del contrato. Cuando dos medicio-

nes independientes, con metodologías distintas y universos distintos, encuentran cero en las mismas variables, lo que hay no es una brecha de implementación. Es una ausencia de instrumento.

El orden de los tres vehículos. Ordenarlos por alcance —documento tipo primero, por ser el de mayor apalancamiento— es lo natural. La medición sugiere reordenarlos por **factibilidad de la primera acción**: la adopción unilateral por resolución de una sola entidad no requiere coordinación, no requiere presupuesto y produce el primer caso publicado, que es lo que hoy no existe. Un documento tipo sin ningún antecedente de aplicación es un documento que se redacta a ciegas; un documento tipo escrito después de que dos o tres entidades hayan aplicado la lista es un documento calibrado. La secuencia recomendada se invierte: **primero el caso, después la norma**.

14. América Latina y la vía de las agencias centralizadas

El atajo institucional que el informe identificó sigue vigente y sigue sin usarse: las agencias de compra centralizadas de la región —Colombia Compra Eficiente, ChileCompra, Perú Compras y sus pares— forman una red natural con foros existentes, y un acuerdo entre dos o tres sobre este estándar o cualquier equivalente crearía el estándar de facto del mercado regional, porque los proveedores preparan su evidencia para el requisito más exigente que enfrentan.

Lo que conviene precisar es **quién está en qué posición**.

País	Instrumento disponible	Estado	Papel en la vía regional
Chile	Bases Tipo para licitación de proyectos de algoritmos e IA con requisitos éticos (enero de 2023); Evaluación de Impacto Algorítmico; medición de sesgo y equidad; Ficha de Transparencia Algorítmica (2023–2025)	En uso, publicados como bienes públicos, con reconocimiento internacional	La pieza probada. No un piloto ni un formulario: el cuerpo de instrumentos más completo de la región
Colombia	CONPES 4144 (106 acciones, 479.273 millones de pesos); estudio diagnóstico de CCE (2026)	Mandato de política sin instrumento de compra	La escala y el mandato; el instrumento falta
Perú	Reglamento de la ley de IA, DS 115-2025-PCM, vigente desde el 22 de enero de 2026, con plazos sectoriales desde el 10 de septiembre de 2026	Norma vigente, aplicación por venir	El primer calendario obligatorio de la región
Brasil	Estándar constitucional de la ADI 6.649 (2022) sobre justificación previa y detallada del uso de datos	Jurisprudencia, no instrumento de compra	El techo jurídico regional

Tabla 10. La región y sus instrumentos, corregida. Fuente: Enpath sobre documentos primarios.

La magnitud de lo que Chile aporta. No es un formulario de transparencia: es un instrumento de licitación completo, publicado dos años antes que este estándar, con evaluación de impacto, medición de sesgo y ficha de transparencia como bienes públicos reutilizables. Colombia aporta la escala y el mandato de su política nacional; el resto adopta por gravedad.

Esto cambia la recomendación regional. Si existe desde enero de 2023 un cuerpo de instrumentos chileno, publicado, aplicado y reconocido internacionalmente, entonces **la vía regional no pasa por construir un estándar nuevo: pasa por adoptar el que ya funciona y complementarlo**. El estándar Enpath aporta lo que las Bases Tipo no cubren —las fases 0, 3 y 4, precisamente las peor cumplidas— y las Bases Tipo aportan lo que el estándar Enpath no tenía: texto de licitación probado en un régimen de contratación real de la región.

La alternativa —treinta países negociando solos contra los mismos cinco proveedores globales— sigue siendo la línea de base. Quince meses después, sigue siendo lo que ocurre.

15. Riesgos, revalorados

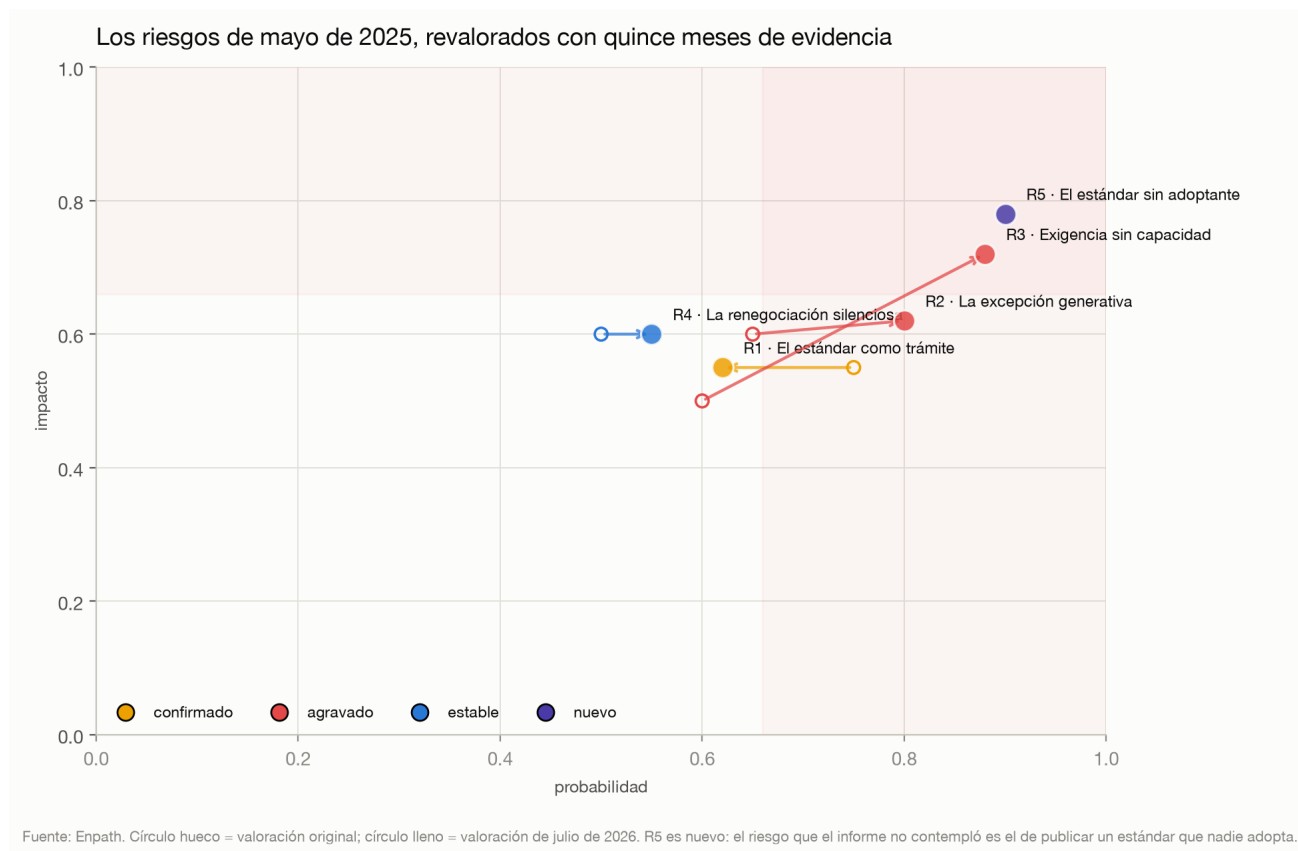


Figura 7. Los riesgos de mayo de 2025 revalorados con quince meses de evidencia. Fuente: Enpath.

Riesgo	Valoración original	Valoración julio de 2026	Evidencia	Mitigación revisada
R1 · El estándar como trámite. Treinta casillas marcadas por el mismo proveedor que redactó el anexo técnico	Probabilidad alta, impacto medio	Confirmado , sin agravamiento	No medible con el scorecard, que evaluó expedientes y no autodeclaraciones	La defensa original se mantiene y se refuerza con la regla de verificación de §16.6
R2 · La excepción generativa. «Es solo un chatbot» como salvoconducto	Probabilidad media-alta, impacto medio	Agravado	La ola generativa creció por canales menores durante todo el periodo; el estándar no llegó a ninguno	Cobertura por objeto y ficha básica universal, reforzadas por el núcleo obligatorio de siete ítems
R3 · Exigencia sin capacidad. El estándar presupone capacidad evaluadora	Probabilidad media, impacto medio	Agravado severamente	Ítem 12 (comité con perfil técnico) 9/20; ítem 21 (gestor técnico) 8/20. La capacidad falta en más de la mitad de los procesos	Comités compartidos entre entidades; el ítem 21 asciende al núcleo obligatorio
R4 · La renegociación silenciosa. Cláusulas que se flexibilizan adición tras adición	Probabilidad media, impacto medio	Estable	Ítem 19 (gestión de cambios de modelo) 2/20: la vigilancia estructural no existe	Publicación de modificaciones en la ficha del registro
R5 · El estándar sin adoptante. Publicar un instrumento que nadie usa	No contemplado	Alto y materializado	Cero adopciones formales localizadas en quince meses	Núcleo obligatorio de siete ítems, resolución modelo del anexo B, y la secuencia invertida de §13

Tabla 11. Riesgos, revalorados. Fuente: Enpath.

Sobre R5. El riesgo que el informe no vio es el que se materializó. El documento se protegió contra el mal cumplimiento —el estándar como trámite, la excepción generativa, la renegociación— y no contra el no-cumplimiento por ausencia de adopción. Es un sesgo característico de los productores de instrumentos: se piensa en cómo se puede desvirtuar el instrumento, no en si alguien lo tomará. La versión 2.0 responde reduciendo el costo de la primera adopción a su mínimo: siete ítems, una resolución de una página y ningún documento nuevo que crear.

Sobre R3. Es el riesgo que más subió y el que menos depende del estándar. Que en poco más de la mitad de los procesos no haya un comité con perfil técnico y en el 60 % no haya gestor de contrato con capacidad técnica significa que la restricción operativa no es normativa sino de personal. Ningún documento resuelve eso. Lo que un documento puede hacer —y la versión 2.0 hace— es **nombrar la restricción** y ofrecer la salida practicable: comités compartidos y acompañamiento de la agencia central mientras la capacidad propia crece.

Parte V · Entrega

16. El estándar versión 2.0

La versión 1.0 se publicó con el compromiso de revisarla con datos. Esta sección entrega la revisión. **Las treinta verificaciones no cambian de contenido.** Cambia lo que las rodea: siete decisiones derivadas de la medición.

16.1 El núcleo obligatorio de siete ítems

El problema práctico de un estándar de treinta ítems con exigencia graduada es que una entidad sin capacidad no sabe por dónde empezar, y el resultado observado es que no empieza. La versión 2.0 designa un **núcleo obligatorio que no admite graduación**: siete verificaciones exigibles a todo sistema comprendido, con independencia de su clasificación, su valor o su canal de compra.

#	Ítem	Fase	Por qué está en el núcleo	Cumplimiento medido
1	Problema y resultados por escrito	0	Sin él no hay objeto que evaluar	7/20 · 35,0 %
2	Clasificación en la matriz de impacto	0	Abre el estándar : sin clasificación, los diez ítems graduados no tienen umbral	0/20 · 0 %
7	Objeto por resultados y métricas, no por marca	1	Sin él el pliego compra una herramienta, no una solución	8/20 · 40,0 %
15	Las diez cláusulas mínimas	2	El núcleo contractual del diagnóstico de agosto de 2024	6/20 · 30,0 %
21	Gestor de contrato con capacidad técnica	3	Da ejecutor a las diez verificaciones de las fases 3 y 4	8/20 · 40,0 %
22	Aceptación contra pruebas pactadas	3	El momento de verdad : ocurre después, con el sistema construido	6/20 · 30,0 %
25	Ficha publicada en el registro	3	Convierte la compra en gobernanza verificable por terceros	0/20 · 0 %

Tabla 12. El núcleo obligatorio de la versión 2.0. Fuente: Enpath.

Los siete promedian **5 de 20**. No es un núcleo elegido por ser fácil: es un núcleo elegido porque sin él los otros veintitrés no son verificables. Dos de los siete están hoy en cero.

16.2 Los cuatro ceros, reasignados a trámites existentes

La lección de §10 es que un requisito sin domicilio institucional no se cumple. La versión 2.0 no crea documentos: **adosa cada uno de los cuatro ceros a un trámite que ya existe y ya tiene firma**.

Ítem en cero	Documento nuevo que exigía la v1.0	Dónde vive en la v2.0
2 · Clasificación en la matriz	Una ficha de clasificación	Un campo obligatorio en los estudios previos , junto al análisis de riesgos que ya se elabora y ya firma el ordenador del gasto
3 · Madurez contrastada	Un contraste de madurez	Un párrafo del análisis del sector , que ya se produce y ya evalúa la capacidad de la entidad
25 · Ficha en el registro	Una ficha de registro	Un requisito del acta de recibo a satisfacción : sin ficha publicada no hay acta, y sin acta no hay pago final
26 · Métricas trimestrales	Una publicación trimestral	Un anexo del informe de supervisión trimestral , que el supervisor ya debe rendir

Tabla 13. Los cuatro ceros, reasignados. Fuente: Enpath.

El cambio del ítem 25 es el más agresivo y el más eficaz: vincular la publicación de la ficha al acta de recibo le da al requisito lo único que le faltaba, que es una consecuencia. Ningún supervisor omite un requisito del que depende un pago.

16.3 El criterio de filtro, sustituido

El hallazgo 3 sale del estándar. Lo reemplaza el criterio derivado de la medición (§11.3): si una entidad solo puede exigir tres cosas, exige **el ítem 2 (clasificar)**, **el ítem 21 (gestor con capacidad técnica)** y **el ítem 22 (aceptación contra pruebas)**.

16.4 Las cláusulas modelo europeas como texto de referencia de la fase 2

Los ítems 15 a 20 remiten, en la versión 2.0, a las **cláusulas modelo de IA de la Comisión Europea en su versión del 5 de marzo de 2025** — versión ligera para sistemas que no son de alto riesgo, versión completa para los que sí, ambas en español. Una entidad ya no tiene que redactar la fase 2: tiene que adaptar un texto existente.

16.5 La Ficha de Transparencia Algorítmica chilena como formato del ítem 25

El ítem 25 exigía «ficha del registro publicada» sin fijar formato, lo que en la práctica significaba que cada entidad debía inventarlo. La versión 2.0 adopta como referencia la **Ficha de Transparencia Algorítmica** publicada por el GobLab de la Universidad Adolfo Ibáñez con ChileCompra y sus socios, compatible con la plantilla de dos niveles del informe de abril de 2025.

16.6 La regla de verificación que faltaba

El estándar v1.0 definía treinta verificaciones y no definía **quién comprueba que la casilla marcada corresponde a un documento que existe**. En una medición externa el problema no aparece; en autoevaluación es el problema central, y el riesgo R1 vive ahí.

Regla de verificación (v2.0). Cada ítem marcado como cumplido debe citar el documento que lo satisface: número de expediente, fecha y ubicación. Un ítem cumplido sin referencia documental se registra como no cumplido. La entidad publica el tablero de sus treinta ítems junto con el expediente del proceso.

Es la misma regla que el estándar le impone al proveedor —la evidencia se ejecuta, no se declara— aplicada por primera vez a la propia entidad.

16.7 La corrección del archivo de datos

El ítem 27 (auditoría externa) figuraba en `data/lista_de_verificacion.csv` con fase «Fase 4 · Ejecución». La fase correcta es «Fase 4 · Operación». Corregido.

17. Metodología y limitaciones

17.1 El instrumento

El estándar fue elaborado con corte al 31 de mayo de 2025, consolidando el diagnóstico y las cláusulas de agosto de 2024, los marcos de la serie 2025 y la convergencia internacional. Cada ítem se justifica por una causa de falla documentada o un requisito internacional convergente. Esa base no se modifica.

17.2 Fuentes y cálculo

Fuente empírica. El archivo `resultados_scorecard.csv` del informe de noviembre de 2025: treinta ítems × veinte procesos, con el número de procesos que cumplen y el número de procesos a los que el ítem era aplicable. 525 pares aplicables de 600 posibles.

Cálculos. Cumplimiento por ítem ($\text{cumplen} \div \text{aplicables}$), por fase ($\text{agregado de cumplen} \div \text{agregado de aplicables}$) y global. La clasificación por productor de la evidencia (§9) se publica completa en el anexo A, para que pueda discutirse ítem por ítem.

Prueba estadística. Permutación de dos colas sobre la diferencia de medias por ítem entre los dos grupos, 200.000 remuestreos, semilla 20250531, implementada en `data/charts_report.py`. Resultado $p = 0,056$. La probabilidad de que los cuatro ceros caigan todos en el grupo de 17 se calcula como $C(17,4)/C(30,4) = 2.380/27.405 = 8,7 \%$.

Verificación documental. Cinco instrumentos internacionales verificados contra fuente primaria en julio de 2026: cláusulas modelo de IA de la UE, AI Playbook británico, memorandos M-25-21 y M-25-22, Bases Tipo de ChileCompra y el estado del *Digital Omnibus*. Búsqueda documental sobre adopciones del estándar en Colombia y la región.

17.3 Limitaciones, declaradas

Limitación	Alcance	Efecto sobre las conclusiones
La muestra del scorecard son veinte procesos seleccionados, no una muestra aleatoria	Todo el análisis empírico	Si la selección privilegió procesos documentados, 27,6 % es un techo. Las comparaciones internas entre ítems y fases son más robustas que el nivel absoluto
La medición es documental	Todo el análisis empírico	Un cumplimiento no documentado cuenta como incumplimiento
Los procesos son anteriores al estándar	Interpretación del 27,6 %	Mide coincidencia espontánea con el consenso internacional, no cumplimiento intentado
La clasificación por productor de la evidencia es una asignación de juicio	§9 y H5	Se publica completa en el anexo A. Reasignaciones razonables de dos o tres ítems no invierten la dirección, pero pueden mover el valor p
n = 30 ítems	La prueba de permutación	Potencia estadística baja. $p = 0,056$ no permite concluir
No se midió tamaño ni origen del proveedor	H6	La afirmación de que el estándar filtra por seriedad y no por tamaño sigue sin prueba
El régimen de contratación asumido es el colombiano	§13, §16	La adaptación a otros regímenes es directa pero no automática

Tabla 14. Limitaciones declaradas. Fuente: Enpath.

17.4 Una lección de método, por tercera vez

Esta serie ha calificado el trabajo chileno por debajo de lo que es en tres informes consecutivos: «piloto» en abril de 2025, «formularios» en mayo de 2025, y en mayo de 2026 se llegó a concluir —erróneamente— que el instrumento no era verificable, conclusión que hubo que revertir. Tres veces el mismo error con tres formas distintas.

La regla de método que se deriva:

Una búsqueda infructuosa no es evidencia de inexistencia. Cuando un instrumento se describe como piloto, formulario, prototipo o experiencia, la calificación debe apoyarse en el documento primario y en su fecha de publicación, no en la impresión que deja una búsqueda. Si el documento primario no se localiza, la afirmación correcta es «no se localizó», no «no existe» ni «es un piloto».

Aplicada aquí, esta regla habría evitado también E3 y E4: dos instrumentos internacionales actualizados en el trimestre previo a la publicación del informe, no recogidos porque se citó la versión conocida sin comprobar si seguía siendo la vigente.

18. Predicciones falsables

Cada predicción tiene umbral, fecha y fuente de verificación. Se declaran fallidas si el umbral no se alcanza.

#	Predicción	Umbral	Fecha	Fuente de verificación
P1	La hipótesis de §9 se confirma: en la próxima medición, los ítems de producción propia se cumplirán menos que los de exigencia al proveedor	Diferencia ≥ 10 pp en el agregado y $p < 0,05$ en permutación con $n \geq 40$ ítems-equivalentes	Scorecard 2027	Archivo de resultados publicado
P2	El cumplimiento global no supera el 35 % sin un instrumento central	Cumplimiento global ≤ 35 % si ninguna agencia central publicó documento tipo antes de la medición	Scorecard 2027	Archivo de resultados publicado
P3	La reasignación de los cuatro ceros a trámites existentes funciona donde se aplica	En las entidades que adopten la v2.0, los ítems 2 y 3 pasan de 0 % a ≥ 60 %	Diciembre de 2027	Estudios previos y análisis del sector publicados en SECOP
P4	El ítem 25 vinculado al acta de recibo es el de mayor efecto marginal	Entre las entidades que adopten la v2.0, el ítem 25 sube más puntos porcentuales que cualquier otro	Diciembre de 2027	Registro público de las entidades adoptantes
P5	Ninguna agencia central de la región publicará un documento tipo de compra de IA	Cero documentos tipo publicados por CCE, ChileCompra o Perú Compras	31 de diciembre de 2027	Sitios institucionales
P6	El retraso europeo se traduce en menos documentación disponible de los proveedores	El ítem 8 (documentación técnica) no supera 70 % en la próxima medición	Scorecard 2027	Archivo de resultados publicado

Tabla 15. Predicciones falsables. Fuente: Enpath.

Sobre P5. Es la predicción más pesimista de esta serie y se publica precisamente por eso. El documento tipo es el instrumento de mayor apalancamiento de toda la agenda de compra, y lo más cercano que existe es un estudio diagnóstico. Si P5 falla —si una agencia central publica un documento tipo antes de que termine 2027—, será la mejor noticia que esta serie pueda recibir, y este informe la registrará como tal.

19. Conclusión

Un estándar que se somete a medición corre el riesgo de que la medición lo contradiga. Este lo corrió y perdió una parte de la apuesta.

Lo que sobrevivió es sustancial. La arquitectura de cinco fases resultó ser el marco correcto para leer el problema: el gradiente 21,1 % – 38,0 % – 35,8 % – 20,2 % – 11,7 % no es ruido, es la forma real de la compra pública de IA en la región, y no se veía antes de que existiera un instrumento que la ordenara así. La graduación por riesgo funcionó: 525 pares aplicables de 600, sin sobrecarga ni insuficiencia. Y las treinta verificaciones se demostraron medibles sobre expedientes reales por un evaluador externo, que es la prueba mínima de que un estándar es un estándar y no una declaración de principios.

Lo que no sobrevivió es la parte que el informe adivinó en lugar de derivar. El hallazgo 3 designó tres ítems como el filtro decisivo por una intuición razonable —los malos proveedores no pueden cumplirlos— y la medición los encontró en el montón medio. Peor: la nota de actualización de este mismo informe leyó los datos como confirmación de lo que ya creía, cuando los datos decían lo contrario. Ese error, cometido sobre el propio archivo publicado, es más grave que la predicción fallida.

Y lo que la medición enseñó no estaba en el informe. **Los cuatro ítems que ningún proceso cumplió no son los más difíciles, ni los más caros, ni los más técnicos: son los cuatro que exigen que alguien dentro de la entidad escriba y firme un documento que antes no tenía casilla.** El ítem 6 —verificar la base legal del tratamiento de datos, que es más difícil que los cuatro— se cumple en 45 %, porque la oficina jurídica ya existía. La diferencia entre 45 % y 0 % es la diferencia entre un requisito con domicilio institucional y uno sin él.

De ahí sale toda la versión 2.0, y de ahí sale la única frase que este informe querría que sobreviviera a su propia corrección: **un requisito que no vive dentro de un trámite existente no se cumple, por barato, sencillo y evidentemente necesario que sea.** No hay que crear el documento. Hay que encontrarle la firma.

Quince meses después de publicar treinta verificaciones, el resultado medible es 27,6 % de cumplimiento espontáneo y cero adopciones formales. El estándar demostró que sirve para medir. Falta demostrar que sirve para comprar. Esa prueba no la da un documento: la da la primera entidad que firme la resolución del anexo B.

Anexos

Anexo A. Los treinta ítems con su medición

Lista completa del estándar, con el resultado del scorecard de noviembre de 2025 y la clasificación por productor de la evidencia introducida en §9. Se publica completa para que la clasificación pueda discutirse ítem por ítem. Ordenada por número de ítem; el archivo fuente es `data/lista_de_verificacion.csv`.

Leyenda de la columna «produce»: E = producción propia de la entidad · P = exigencia al proveedor. **Núcleo** marca los siete ítems del núcleo obligatorio de la versión 2.0.

#	Verificación	Fase	Aplica a	Cumple	Tasa	Produce	Núcleo
1	Problema y resultados esperados definidos por escrito	0	Todos	7/20	35,0 %	E	●
2	Clasificación del sistema en la matriz impacto × reversibilidad	0	Todos	0/20	0 %	E	●
3	Nivel de madurez contrastado con la regla de ambición	0	Todos	0/20	0 %	E	
4	Alternativas no algorítmicas evaluadas y documentadas	0	Todos	3/20	15,0 %	E	
5	Diálogo de mercado sobre el problema, no sobre una solución	0	Medio y alto	5/14	35,7 %	E	
6	Base legal del tratamiento de datos verificada	0	Todos	9/20	45,0 %	E	
7	El objeto describe resultados y métricas; no una marca ni una técnica	1	Todos	8/20	40,0 %	P	●
8	Documentación técnica del sistema, con alcance según riesgo	1	Todos	12/20	60,0 %	P	
9	Ficha de datos de entrenamiento y operación	1	Medio y alto	2/14	14,3 %	P	
10	Desempeño desagregado por grupos relevantes	1	Alto impacto sobre personas	1/9	11,1 %	P	
11	Prueba práctica sobre datos y casos de la entidad, como criterio	1	Medio y alto	4/14	28,6 %	P	
12	Criterios de evaluación con pesos publicados; comité con perfil técnico	1	Todos	9/20	45,0 %	E	
13	Requisitos de seguridad y continuidad (riesgo de terceros)	1	Todos	13/20	65,0 %	P	
14	Costo total de la vida útil exigido en la oferta	1	Todos	3/20	15,0 %	P	
15	Las diez cláusulas mínimas de agosto de 2024 incorporadas	2	Todos	6/20	30,0 %	P	●
16	Niveles de servicio sobre métricas de desempeño y disponibilidad	2	Todos	14/20	70,0 %	P	
17	Cronograma de entregas con hitos de validación verificables	2	Todos	10/20	50,0 %	P	
18	Propiedad de datos y artefactos; exportación pactada con formato	2	Todos	7/20	35,0 %	P	
19	Gestión de cambios de modelo con aprobación previa de la entidad	2	Todos	2/20	10,0 %	P	
20	Plan de salida con periodo de transición asistida	2	Todos	4/20	20,0 %	P	
21	Gestor de contrato con capacidad técnica designado	3	Todos	8/20	40,0 %	E	●
22	Aceptación contra pruebas pactadas, nunca contra demostración	3	Todos	6/20	30,0 %	E	●

#	Verificación	Fase	Aplica a	Cumple	Tasa	Produce	Núcleo
23	Protocolo de supervisión aprobado antes del despliegue	3	Medio y alto	3/14	21,4 %	E	
24	Línea de base de desempeño medida antes de operar	3	Todos	2/20	10,0 %	E	
25	Ficha del registro de algoritmos publicada	3	Todos	0/20	0 %	E	•
26	Métricas trimestrales reportadas y publicadas	4	Todos	0/17	0 %	E	
27	Auditoría externa según calendario pactado	4	Medio y alto	1/12	8,3 %	E	
28	Monitoreo de deriva con umbrales y responsable	4	Medio y alto	1/14	7,1 %	E	
29	Renovación solo por puerta 4, con evaluación documentada	4	Todos	3/11	27,3 %	E	
30	Al terminar: exportación verificada y cierre documentado	4	Todos	2/6	33,3 %	E	
Total				145/525	27,6 %	17 E · 13 P	7

Tabla 16. El estándar completo, medido. Fuente: Enpath; scorecard de noviembre de 2025, archivo resultados_scorecard.csv. El ítem 27 figura en el archivo original con fase «Fase 4 · Ejecución»; corregido aquí a «Fase 4 · Operación».

A.1 Notas sobre la clasificación por productor

Tres asignaciones admiten discusión legítima y se declaran:

Ítem 12 (criterios con perfil técnico en el comité) → E. Podría argumentarse que es una exigencia de pliego. Se clasifica como producción propia porque lo que se verifica no es el párrafo del pliego sino la composición efectiva del comité, que la entidad debe integrar con personas que tiene o no tiene.

Ítem 15 (las diez cláusulas mínimas) → P. Es un texto que la entidad redacta, pero lo que se verifica es que el contrato firmado las contenga, lo que depende de que el proveedor las acepte. Se clasifica como exigencia al proveedor. Reasignarlo a E movería la media de E hacia arriba y la de P hacia abajo, reduciendo la diferencia — es la reasignación individual más adversa a la hipótesis de \$9, y se señala por eso.

Ítem 20 (plan de salida) → P. Mismo razonamiento que el 15.

Anexo B. Resolución modelo para adoptar el estándar

Texto adaptable; no es un concepto jurídico. Diseñado para ocupar una página y no requerir presupuesto, coordinación ni norma superior — las tres razones por las que la adopción no ocurrió en quince meses.

RESOLUCIÓN [número] DE [fecha] Por la cual se adopta el estándar de verificación para la adquisición de sistemas de inteligencia artificial y de decisión automatizada de [entidad]

- 1. Objeto y alcance.** Adóptase el estándar de treinta verificaciones en cinco fases para la adquisición de sistemas de inteligencia artificial y de decisión automatizada. Se aplica **por objeto**: a todo proceso cuyo resultado sea un sistema que decida, apoye o recomiende decisiones sobre personas o recursos públicos, **con independencia del canal de compra, de su valor y de su denominación técnica**.
- 2. Núcleo obligatorio.** Los ítems 1, 2, 7, 15, 21, 22 y 25 son exigibles a todo proceso comprendido en el numeral 1, sin graduación por riesgo.
- 3. Exigencia graduada.** Los demás ítems se aplican según la clasificación del sistema en la matriz de impacto × reversibilidad, conforme al anexo A.
- 4. Integración con los trámites existentes.** No se crean documentos nuevos: a) La clasificación en la matriz (ítem 2) y el contraste de madurez (ítem 3) se incorporan como apartados obligatorios de **los estudios previos y del análisis del sector**. b) La publicación de la ficha en el registro (ítem 25) es **requisito del acta de recibo a satisfacción**. Sin ficha publicada no se suscribe el acta. c) Las métricas trimestrales (ítem 26) se incorporan como anexo del **informe de supervisión trimestral**.
- 5. Fase 2.** Para los ítems 15 a 20, la dependencia jurídica tomará como texto de referencia las **cláusulas modelo de contratación de inteligencia artificial de la Comisión Europea en su versión del 5 de marzo de 2025**, en su versión completa o ligera según la clasificación del sistema, adaptadas al régimen de contratación aplicable.
- 6. Regla de verificación.** Cada ítem marcado como cumplido citará el documento que lo satisface: expediente, fecha y ubicación. **Un ítem cumplido sin referencia documental se registra como no cumplido**. El tablero de los treinta ítems se publica junto con el expediente del proceso.
- 7. Capacidad.** Cuando la entidad no cuente con perfil técnico para integrar el comité (ítem 12) o designar el gestor (ítem 21), solicitará acompañamiento a [agencia central / entidad par] y dejará constancia en el expediente. **La ausencia de capacidad se documenta; no exime**.
- 8. Régimen de transición.** Los contratos vigentes a la fecha de esta resolución aplicarán los ítems de las fases 3 y 4 desde su siguiente informe de supervisión, y los ítems 25 y 26 dentro de los [180] días siguientes.
- 9. Responsable.** [Cargo] responde por la aplicación del estándar y publica semestralmente el consolidado de cumplimiento de la entidad.

Notas de diseño. El numeral 4 es el corazón de la versión 2.0: retira los cuatro ceros del limbo y los deposita en trámites que ya existen, ya tienen dueño y ya tienen firma. El numeral 6 cierra el riesgo R1 aplicándole a la entidad la misma regla que le impone al proveedor. Y el numeral 7 hace lo único honesto que puede hacerse con el riesgo R3: reconocer que la capacidad falta, obligar a decirlo por escrito y no aceptarlo como excusa.

Anexo C. Los siete cambios de la versión 2.0

#	Cambio	Origen en la evidencia	Sección
1	Núcleo obligatorio de siete ítems sin graduación	Cero adopciones en 15 meses: el costo de entrada era demasiado alto	\$16.1
2	Los cuatro ceros reasignados a trámites existentes	Los cuatro ceros son los cuatro ítems sin domicilio institucional	\$16.2
3	Criterio de filtro sustituido: ítems 2, 21 y 22	El hallazgo 3 original quedó en los puestos 10, 13 y 16 de 30	\$11.3, \$16.3
4	Cláusulas modelo de la UE (5 de marzo de 2025) como texto de la fase 2	Existe texto contractual modelo oficial en español, con versión ligera	\$12.1, \$16.4
5	Ficha de Transparencia Algorítmica chilena como formato del ítem 25	El ítem 25 no fijaba formato y quedó en 0/20	\$16.5
6	Regla de verificación documental de cada casilla	El estándar no decía quién comprueba lo marcado; ahí vive el riesgo R1	\$16.6
7	Corrección de la fase del ítem 27 en el archivo de datos	Error de transcripción que rompe la agregación por fase	\$16.7

Tabla 17. Los siete cambios de la versión 2.0. Fuente: Enpath.

Anexo D. Glosario

Término	Definición operativa en este documento
Verificación	Un ítem del estándar: una comprobación binaria con criterio de cumplimiento y ámbito de aplicación declarado
Aplicable	Par ítem-proceso en el que el ítem correspondía al proceso según la graduación por riesgo. 525 de 600 posibles
Graduación por riesgo	Regla que hace depender la exigencia de la matriz de impacto × reversibilidad, no del valor del contrato
Núcleo obligatorio	Los siete ítems de la versión 2.0 exigibles a todo sistema sin graduación
Producción propia	Ítem cuya evidencia debe generar la entidad: 17 de 30
Exigencia al proveedor	Ítem cuya evidencia produce el proveedor a solicitud de la entidad: 13 de 30
Domicilio institucional	Trámite existente, con dueño y firma, donde un requisito puede depositarse. Su ausencia explica los cuatro ceros
Cobertura por objeto	Regla de alcance: el estándar aplica a lo que el sistema decide, no al canal por el que se compró
Coherencia retroactiva	Leer un dato nuevo como confirmación de lo que ya se creía. El error de la

Referencias

Instrumentos internacionales de compra de IA

1. Comisión Europea, Public Buyers Community (2025). *Updated EU AI model contractual clauses*, 5 de marzo de 2025: versión completa para alto riesgo, versión ligera y comentario, en 24 lenguas. Actualiza la versión del 29 de septiembre de 2023. <https://public-buyers-community.ec.europa.eu/communities/procurement-ai/resources/updated-eu-ai-model-contractual-clauses>
2. Government Digital Service, Reino Unido (2025). *Artificial Intelligence Playbook for the UK Government*, 10 de febrero de 2025. <https://www.gov.uk/government/publications/ai-playbook-for-the-uk-government>
3. Gobierno del Reino Unido (2020). *Guidelines for AI procurement* (superada por el AI Playbook de febrero de 2025). <https://www.gov.uk/government/publications/guidelines-for-ai-procurement>
4. Office of Management and Budget, Estados Unidos (2025). *Memorandos M-25-21 y M-25-22*, 3 de abril de 2025 (uso y adquisición de IA en agencias federales). <https://www.whitehouse.gov/omb/information-for-agencies/memoranda/>
5. World Economic Forum (2020). *AI Procurement in a Box*. <https://www.weforum.org/publications/ai-procurement-in-a-box/>
6. Unión Europea (2024). *Reglamento (UE) 2024/1689*. <https://eur-lex.europa.eu/eli/reg/2024/1689/oj> · (2026) *Digital Omnibus on AI*.

Instrumentos regionales

1. ChileCompra (2023). *Bases Tipo para la licitación de proyectos de algoritmos e inteligencia artificial con requisitos éticos*, enero de 2023. <https://www.chilecompra.cl/2023/01/ya-se-encuentra-disponible-bases-tipo-para-licitar-proyectos-de-algoritmos-e-inteligencia-artificial-con-requisitos-eticos/>
2. GobLab, Universidad Adolfo Ibáñez. *Algoritmos Éticos, Responsables y Transparentes: Evaluación de Impacto Algorítmico, medición de sesgo y equidad, y Ficha de Transparencia Algorítmica*. <https://goblab.uai.cl/>
3. Agencia Nacional de Contratación Pública – Colombia Compra Eficiente (2026). *Radar de la compra pública*, boletín edición N° 3 de 2026: estudio diagnóstico sobre la contratación pública de soluciones tecnológicas que incorporan inteligencia artificial. <https://www.colombiacompra.gov.co/archivos/boletin/boletin-edicion-n-3-de-2026-radar-de-la-compra-publica>
4. Departamento Nacional de Planeación (2025). *Documento CONPES 4144*. <https://colaboracion.dnp.gov.co/CDT/Conpes/Econ%C3%B3micos/4144.pdf>
5. Presidencia del Consejo de Ministros del Perú (2025). *Decreto Supremo 115-2025-PCM*, reglamento de la ley de inteligencia artificial; vigente desde el 22 de enero de 2026.

Serie de investigación de Enpath

1. Enpath (2024). *Por qué los gobiernos compran mal la IA: ocho causas y diez cláusulas* (agosto); *catálogo de fallas* (junio); *marco de confianza institucional* (octubre). Esta biblioteca.
2. Enpath (2025). *Marco de gobernanza de IA para ministerios* (enero); *modelo de madurez* (febrero); *marco de gobernanza de datos gubernamentales* (marzo); *registros de algoritmos* (abril). Esta biblioteca.
3. Enpath (2025). *Scorecard de compra pública de IA* (noviembre), archivo resultados_scorecard.csv — **la medición empírica de este estándar sobre veinte adquisiciones reales**. Esta biblioteca.
4. Enpath (2026). *Índice Enpath de gobernanza de IA del sector público: Colombia 2026* (marzo), archivo `igia_2026.csv`; *Marco Enpath de compra responsable de IA* (enero). Esta biblioteca.

Notas

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

1. Comisión Europea, Public Buyers Community. Las cláusulas modelo de contratación de IA (MCC-AI) fueron actualizadas el **5 de marzo de 2025** y publicadas en tres documentos: versión completa para IA de alto riesgo, alineada con las obligaciones del capítulo III del Reglamento (UE) 2024/1689 adoptado el 13 de junio de 2024; versión ligera, personalizable, para sistemas que no son de alto riesgo; y un comentario con orientación de uso, personalización y aplicación. Los tres textos están disponibles en veinticuatro lenguas. La versión anterior databa del 29 de septiembre de 2023 y precedía a la adopción del Reglamento. Es el texto de referencia de la fase 2 en la versión 2.0 del estándar (§16.4).
2. *Artificial Intelligence Playbook for the UK Government*, publicado por el Government Digital Service el **10 de febrero de 2025**. Reemplaza y amplía el *Generative AI Framework for HMG* de enero de 2024. Establece diez principios con controles prácticos y orientación de contratación y aseguramiento para la selección, compra, despliegue y gobierno de sistemas de IA por el gobierno central y los organismos en su alcance. Supera a las *Guidelines for AI procurement* de 2020 y cubre las fases 3 y 4 del estándar, que son las peor cumplidas.
3. *Digital Omnibus on AI* de la Unión Europea: adoptado 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; el deber de transparencia del artículo 50 mantiene el 2 de agosto de 2026.
4. Agencia Nacional de Contratación Pública – Colombia Compra Eficiente, boletín *Radar de la compra pública*, edición N° 3 de 2026. Anuncia un estudio diagnóstico sobre la contratación pública de soluciones tecnológicas que incorporan inteligencia artificial. No se localizó, a 25 de julio de 2026, ningún documento tipo, guía ni condición de entrada en instrumentos agregados que incorpore requisitos de compra de IA.
5. Índice Enpath de gobernanza de IA del sector público, Colombia 2026 (marzo de 2026), archivo publicado `igia_2026.csv`: dieciocho ministerios; acto de gobernanza en 4, inventario publicado en 2, ficha pública de sistema en 2; cero en métricas publicadas, evaluación de riesgo publicada y vía de impugnación específica. Los ceros del IGIA y los del scorecard coinciden en las mismas variables por dos metodologías independientes.