

ENPATH

INFORME · INVESTIGACIÓN

Marco Enpath de Compra Responsable de IA

El estándar: cinco principios, treinta ítems y la conformidad demostrable

Enero de 2026

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

Enpath · Investigación · El Estándar Enpath · Pilar: Compra pública Publicado: enero de 2026 · **Cor- te de la evidencia:** 24 de julio de 2026 **Versión del estándar:** v2 · **Ámbito:** Colombia y América Lati- na, con referencias comparadas **Clasificación de la evidencia:** estándar propio calibrado sobre me- didión empírica propia y fuentes normativas primarias (ver §18.3)

Dos años de trabajo sobre compra pública de IA — el diagnóstico (2024), el estándar operativo (2025) y la primera medición empírica de la región (2025) — se consolidan aquí en una sola pieza con nombre propio: el Marco Enpath de compra responsable de IA. Cinco principios que no cambian, treinta ítems calibrados por la evidencia y, la novedad que la evidencia exigió: **cómo se demuestra que se adoptó.**

Tabla de contenido

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

Parte II — Contexto 4. Antecedentes: la trayectoria de la agenda de compra 5. El problema: por qué falla la compra pública de IA 6. Marco normativo aplicable

Parte III — El estándar 7. Los cinco principios 8. El estándar v2: tres condiciones de entrada, regla de canal, graduación 9. La conformidad: declaración, autoevaluación y verificación externa 10. La gobernanza del marco

Parte IV — Evidencia 11. Análisis de datos: los treinta ítems, medidos 12. Comparación internacio- nal y estudios de caso 13. Análisis económico y de actores

Parte V — Prospectiva y decisión 14. Riesgos 15. Oportunidades 16. Recomendaciones y las tres ru- tas de adopción 17. Perspectivas y predicciones falsables 18. Metodología, calidad de la evidencia y limitaciones 19. Conclusión

Anexos - Anexo A. Texto modelo de la declaración de conformidad - Anexo B. Clasificación de los treinta ítems (G1 / G2) y datos del análisis - Anexo C. Criterios de verificación de las tres condicio- nes de entrada - Anexo D. Protocolo de verificación externa por muestreo - Anexo E. Glosario - Anexo F. Índice de figuras y tablas - Referencias

Parte I — Síntesis

1. Resumen ejecutivo

Un estándar merece llamarse estándar cuando cumple tres condiciones: **contenido estable, evi- dencia de calibración y forma de demostrar su adopción.** La agenda de compra de esta serie cum- plió las dos primeras por etapas —las ocho causas y las diez cláusulas (agosto de 2024), los treinta ítems en cinco fases (mayo de 2025), la línea de base empírica (noviembre de 2025)— y el

documento de enero de 2026 añadió la tercera. El **Marco Enpath de compra responsable de IA (MCR)** integra las piezas en tres niveles.

Los cinco principios — el contenido que no cambia entre versiones: el problema antes que la tecnología; la exigencia sigue al riesgo, no al canal ni al precio; nada sin dueño, tampoco después de la firma; los datos y la salida pertenecen al comprador; todo verificable y público. Cualquier revisión futura que contradiga un principio no es una versión: es otro marco.

El estándar v2 — los treinta ítems de mayo de 2025 con las tres calibraciones que la medición exigió: las **condiciones de entrada** (clasificación de riesgo, ficha en el registro y métricas de operación dejan de ser ítems entre treinta y pasan a ser requisitos sin los cuales el proceso no se considera conforme); la **regla de canal** (el estándar aplica por objeto —sistema que decide o apoya decisiones—, cerrando la vía de la contratación menor); y la **graduación simplificada** para bajo impacto (doce ítems en lugar de treinta, para que la exigencia no empuje a la informalidad).

La conformidad — la pieza nueva: una **declaración de conformidad** de una página por proceso, publicada con el expediente (Anexo A); la **autoevaluación anual** de la entidad con el instrumento del scorecard; y la **verificación externa por muestreo** (Anexo D) como el mecanismo que separa la adopción real de la retórica. La regla de honestidad del marco: **adoptar el MCR es publicar la declaración; lo demás es citarlo.**

1.1. La línea de base, con su aritmética verificada

La calibración del estándar se apoya en el archivo publicado del scorecard de noviembre de 2025, recomputado íntegro:

Indicador	Valor verificado
Ítems evaluados	30, en cinco fases
Procesos evaluados	20 (denominadores por ítem entre 6 y 20, por graduación de riesgo)
Mediana de cumplimiento por ítem	27,9 %
Media de cumplimiento por ítem	26,7 %
Mejor ítem	Niveles de servicio sobre métricas: 14 de 20 (70 %)
Ítems en cero absoluto	Cuatro
Cumplimiento medio de la fase 0 (preparación)	21,8 %
Cumplimiento medio de la fase 2 (contrato)	35,8 % — el máximo
Cumplimiento medio de la fase 4 (operación y cierre)	15,2 % — el mínimo

Tabla 1. La línea de base del estándar, recomputada. Fuente: scorecard Enpath de compra pública de IA (noviembre de 2025), archivo resultados_scorecard.csv.

1.2. El hallazgo nuevo

El documento de enero afirmó que los estándares de compra «funcionan por formularios, no por convicciones». Esta edición lo prueba. Clasificando los treinta ítems en dos grupos —**G1**, requisitos que ya existen en los formatos habituales de contratación tecnológica (niveles de servicio, seguridad, documentación técnica, hitos, aceptación, gestor de contrato); **G2**, requisitos específi-

cos de sistemas algorítmicos (clasificación de riesgo, ficha de datos, desempeño por grupos, gestión de cambios de modelo, registro, métricas, auditoría, deriva)— el resultado es el siguiente:

- **G1 (13 ítems): 43,5 % de cumplimiento medio, mediana 40 %, cero ítems en cero.**
- **G2 (17 ítems): 13,9 % de cumplimiento medio, mediana 11,1 %, los cuatro ceros absolutos.**
- **Diferencia de medias: 29,6 puntos porcentuales, con $p \approx 0,00001$ en una prueba de permutación de 200.000 remuestreos.**

La implicación de diseño es exactamente la que el marco v2 aplicó, y ahora tiene fundamento cuantitativo: **convertir en formato obligatorio lo que la convicción no logró en dos años.** Y la implicación de política es más incómoda: la brecha no mide falta de capacidad ni de voluntad de las entidades — mide la ausencia de un formulario donde escribir el requisito. La clasificación completa está en el Anexo B para que cualquiera pueda disputarla y recalcular.

1.3. Dos precisiones sobre la línea de base

Dos correcciones importan para el propio diseño del marco.

La primera es aritmética. El original habla de «los tres ceros estructurales de la medición» y construye sobre ellos las tres condiciones de entrada. La medición encontró **cuatro** ítems en cero: clasificación en matriz de impacto (0/20), **madurez contrastada o regla de ambición** (0/20), ficha publicada en registro (0/20) y métricas trimestrales publicadas (0/17). El cuarto no fue elevado a condición de entrada y el documento no explicaba por qué. Esta edición propone la explicación y la decisión.

La segunda es de mapa regional, y es la más consecuente. El original presenta la «ruta de la agencia central» como la apuesta de 2026 y menciona la práctica chilena solo como referencia de transparencia. La verificación muestra que **ChileCompra publicó Bases Tipo para la licitación de proyectos de algoritmos e inteligencia artificial con requisitos éticos, disponibles desde enero de 2023**, tras consulta pública entre octubre y noviembre de 2022, con hitos de transparencia, privacidad, no discriminación y explicabilidad, y con exigencias de métricas estadísticas de equidad y análisis de sesgo según la naturaleza del contrato; el proyecto que las produjo —ejecutado por el GobLab de la Universidad Adolfo Ibáñez con apoyo de BID Lab desde 2020— entregó además una evaluación de impacto algorítmico, una herramienta de medición de sesgo y una ficha de transparencia algorítmica como bienes públicos¹.

La consecuencia no es que el MCR sobre. Es que **su casilla vacía queda a la vista** (Figura 7): de los cinco estándares comparados, el MCR es el único con conformidad publicada y con instrumento de medición propio, y es también **el único que ninguna agencia central ha adoptado**. La apuesta de 2026 no era sobre viabilidad —Chile la demostró— sino sobre voluntad.

2. Hallazgos principales

H1. La evidencia mandó sobre el diseño. Las tres modificaciones de la v2 no salen de deliberación editorial sino de la medición: donde el scorecard encontró ceros, el marco pone condiciones de entrada; donde encontró el peor cumplimiento, pone la regla de cobertura; donde el costo de exigencia podía empujar a la informalidad, gradúa. Esta edición añade la prueba de que el mecanismo supuesto existe (§11.3). *Confianza: alta.*

H2. Los estándares de compra se cumplen donde ya hay formato — y ahora es un resultado, no una intuición. Los requisitos genéricos de contratación tecnológica cumplen el 43,5 %; los específicos de sistemas algorítmicos, el 13,9 %; los cuatro ceros están todos en el segundo grupo; la diferencia no es atribuible al azar ($p \approx 0,00001$). **La consecuencia de diseño es que un estándar de IA solo funciona si se convierte en campo de un formulario existente.** *Confianza: alta (cálculo sobre datos propios publicados; la clasificación de ítems es declarada y disputable).*

H3. El cumplimiento no decae de forma gradual: cae donde la obligación deja de existir. El máximo está en la fase de contrato (35,8 %), que es la fase con formatos establecidos, y el mínimo en la fase de operación y cierre (15,2 %), donde ninguna norma exige nada. La fase 0 es la segunda peor (21,8 %) por la misma razón: nadie tiene un formulario para clasificar riesgo. *Confianza: alta.*

H4. La conformidad declarada sigue siendo el eslabón que falta en la región — incluido en el país que más avanzó. Chile tiene Bases Tipo adoptadas por su agencia central desde enero de 2023, con requisitos éticos exigibles en el pliego, y no tiene declaración de conformidad publicada por proceso. La contribución específica del MCR se reduce y se afila: no es «el estándar de compra de IA de la región», es el que añade la demostración pública de adopción y el instrumento para medirla. *Confianza: alta.*

H5. La medición encontró cuatro ceros y el marco atendió tres. El cuarto —la regla de ambición sobre madurez del proveedor— quedó sin tratamiento diferenciado. Esta edición lo resuelve trasladándolo a campo obligatorio de justificación en la declaración de conformidad, con el argumento de verificabilidad que la §8.2 expone. *Confianza: alta sobre el dato; la decisión de diseño es propia y discutible.*

H6. Las tres rutas de adopción tienen costos y alcances distintos, son acumulables, y la segunda ya no es hipotética. La ruta de entidad (resolución, semanas) saca a una organización de la banda del 28 %; la de agencia central (documentos tipo, meses) mueve la mediana nacional —y Chile demostró en 2023 que es alcanzable—; la regional (acuerdo de agencias y condiciones de banca, años) fija el estándar de facto del mercado. El error sería esperar la tercera para empezar la primera. *Confianza: media-alta.*

H7. La asimetría documental del mercado se mantiene, con dieciséis meses más de plazo. La industria global producirá documentación técnica, gestión de riesgos, supervisión humana y registro por obligación europea, con vista al 2 de diciembre de 2027 en lugar del 2 de agosto de 2026. **Exigir el estándar sigue siendo barato y no exigirlo sigue dejando valor sobre la mesa** — con el riesgo nuevo de que el aplazamiento se invoque como razón para esperar. *Confianza: alta.*

H8. La graduación por riesgo ya operaba antes de escribirse. Diez de los treinta ítems del scorecard tienen denominadores menores a veinte, es decir, no aplicaban a todos los procesos. La v2 no introdujo la graduación: la formalizó. Es un argumento a favor de su viabilidad práctica. *Confianza: alta.*

H9. 2026 era el año de la prueba y la prueba se está perdiendo por la casilla vacía. La apuesta de noviembre de 2025 —si al menos una agencia central adopta el estándar en 2026, la mediana de su subgrupo superará el 45 % en la siguiente medición— sigue en pie y, a julio de 2026, sin agencia central colombiana que la haya tomado. La agencia nacional de contratación ha comenzado a producir orientación sobre compra de soluciones con IA en sus boletines de 2026, pero no se identificó un documento tipo específico. *Confianza: media (verificación documental pública, no exhaustiva).*

3. Por qué importa

La compra sigue siendo la política de IA que decide más sistemas por año en la región — y desde noviembre de 2025, la única con línea de base numérica.

Este documento apoya la decisión que 2026 puso sobre tres mesas. La de cada **entidad**: adoptar por resolución y publicar la primera declaración de conformidad, con el texto del Anexo A. La de las **agencias centrales**: documentos tipo con el estándar, la intervención de mayor apalancamiento medible de toda la agenda — y con un precedente regional verificado que retira la excusa de la inviabilidad. La de la **banca multilateral**: condicionar financiamiento digital a la conformidad, la ruta que convierte el estándar en regional sin esperar tratados, y que en el caso chileno ya financió la producción del instrumento.

El costo de cada ruta está tasado; el costo de ninguna, también: **la mediana del 27,9 % es el precio anual del statu quo**, y su composición es ahora conocida con precisión de ítem. Los contratos que se firmen en los próximos doce meses tendrán vigencias de años, y cada uno firmado sin las tres condiciones de entrada es un sistema que operará sin clasificación de riesgo, sin ficha pública y sin métricas durante toda su vida útil.

Parte II — Contexto

4. Antecedentes: la trayectoria de la agenda de compra

Fecha	Pieza	Qué aportó	Qué dejó pendiente
Agosto de 2024	Las ocho causas de la mala compra pública de IA y las diez cláusulas mínimas	El diagnóstico y la respuesta contractual mínima	El estándar completo del proceso
Mayo de 2025	Estándar de compra pública de IA: treinta ítems en cinco fases, con graduación por riesgo	El contenido operativo	La medición
Noviembre de 2025	Scorecard: primera medición empírica regional, veinte procesos	La línea de base y la calibración	La forma de demostrar adopción
Enero de 2026	MCR: principios, estándar v2 y conformidad	La demostrabilidad y la gobernanza del estándar	Su primera adopción real
Mayo de 2026	Estándares de transparencia: las cinco ventanas	Cierra el flanco postcontractual que el scorecard dejó más expuesto	—

Tabla 2. La trayectoria de la agenda de compra. Fuente: Enpath, serie de investigación 2024–2026.

El MCR es la consolidación con nombre, versión y gobernanza — el formato que permite a terceros adoptarlo, citarlo y auditarlo sin leer tres años de serie. Y su relación con el resto de la caja de herramientas es deliberada: **el MCR es a la compra lo que el marco ministerial (enero de 2025) es a la gobernanza y la evaluación específica de riesgo (septiembre de 2025) al riesgo**. Los tres comparten matriz de impacto, puertas, registro y métricas: una entidad que adopta uno tiene aproximadamente el 60 % del siguiente. Los instrumentos se compraron por separado; gobiernan juntos.

5. El problema: por qué falla la compra pública de IA

5.1. Las ocho causas, revisitadas con la medición

El diagnóstico de agosto de 2024 identificó ocho causas. La medición de noviembre de 2025 permite ahora ordenarlas por su huella empírica:

Causa (2024)	Huella en la medición (2025)
Ausencia de problema definido antes de la tecnología	Problema y resultados por escrito: 7 de 20 (35 %); alternativas no algorítmicas evaluadas: 3 de 20 (15 %)
Exigencia atada al precio y al canal, no al riesgo	Clasificación en matriz de impacto: 0 de 20
Evaluación sin capacidad técnica	Criterios con perfil técnico en comité: 9 de 20 (45 %); prueba práctica como criterio: 4 de 14 (29 %)
Contratos sin las cláusulas mínimas	Diez cláusulas mínimas: 6 de 20 (30 %); gestión de cambios de modelo: 2 de 20 (10 %)
Dependencia de proveedor no pactada	Propiedad y exportación de datos: 7 de 20 (35 %); plan de salida: 4 de 20 (20 %)
Operación sin dueño ni métricas	Métricas trimestrales publicadas: 0 de 17 ; monitoreo de deriva: 1 de 14 (7 %)
Secreto comercial contra transparencia pública	Reservas genéricas documentadas; ficha publicada en registro: 0 de 20
Renovación por inercia	Renovación con evaluación documentada: 3 de 11 (27 %); cierre con exportación verificada: 2 de 6 (33 %)

Tabla 2. Las ocho causas y su huella medida. Fuente: Enpath, cruce del diagnóstico de agosto de 2024 con el scorecard de noviembre de 2025.

La tabla es el argumento de que el diagnóstico de 2024 no era retórico: **cada causa tiene su número, y los números de las causas más graves son ceros o cifras de un dígito.**

5.2. Por qué la compra es el punto de intervención correcto

Tres razones, en orden de fuerza.

Porque es donde el Estado decide qué sistemas existirán. Una política nacional orienta; un contrato crea. El sistema que se compra sin clasificación de riesgo operará años sin ella, con independencia de lo que cualquier política diga después.

Porque el momento de máximo poder de negociación de la entidad es antes de firmar. Después de la firma, exigir documentación técnica, métricas o exportación de datos es renegociar. Antes, es un requisito del pliego que el proveedor incorpora en su oferta sin costo marginal relevante — y esa asimetría temporal es el activo que la mediana del 27,9 % está desperdiciando.

Porque la compra es medible desde fuera. El expediente es público por regla general. Es la única capa de toda la agenda de gobernanza de IA que un tercero puede auditar sin acceso privilegiado — y por eso el scorecard existe y el índice de gobernanza de marzo de 2026 pudo construirse.

5.3. Por qué un estándar y no una norma

El marco existe como estándar abierto, no como propuesta normativa, por tres razones prácticas: no requiere aprobación legislativa; puede adoptarse en tres rutas de costo y plazo distintos (§16); y puede corregirse con evidencia cada año, cosa que una norma no puede. La contrapartida es la que la §14 trata como riesgo R5: **un estándar sin adoptante no es nada, y a julio de 2026 el MCR no tiene ninguna agencia central que lo haya adoptado.**

6. Marco normativo aplicable

Instrumento	Qué aporta a la compra de IA	¿Exige alguna de las tres condiciones de entrada?
Régimen colombiano de contratación estatal	Publicidad del expediente, selección objetiva, supervisión del contrato	No, para el objeto IA
Ley 1712 de 2014 (transparencia y acceso)	Publicidad como regla; reserva excepcional, taxativa, motivada y temporal	Habilita la publicación de la declaración de conformidad; no la exige
Ley 1581 de 2012 y Ley 1266 de 2008 (datos personales)	Base legal del tratamiento; principios de finalidad y proporcionalidad	Indirectamente: la base legal de datos es ítem del estándar (9 de 20 en la medición)
Circular Externa 002 del 21 de agosto de 2024 (autoridad de datos)	Lineamientos sobre tratamiento de datos personales en sistemas de IA: idoneidad, necesidad, razonabilidad y proporcionalidad en sentido estricto; recomendación de estudios de impacto de privacidad ⁶	Fundamenta la clasificación de riesgo como exigencia: un sistema cuyo tratamiento no supera el test no debería contratarse. No impone su forma
CONPES 4144 (14 de febrero de 2025)	Política nacional de IA: seis ejes, más de cien acciones a 2030, 479.000 millones de pesos ⁷	No: la compra es su instrumento más concreto y menos ejecutado
Documentos tipo de la agencia nacional de contratación	El vehículo de mayor alcance disponible	No, todavía: es la decisión pendiente de mayor retorno del bienio
Reglamento (UE) 2024/1689 y cláusulas modelo europeas	El contenido del estándar que la oferta global producirá	Exigible por contrato desde el próximo pliego, sin esperar fechas europeas ²

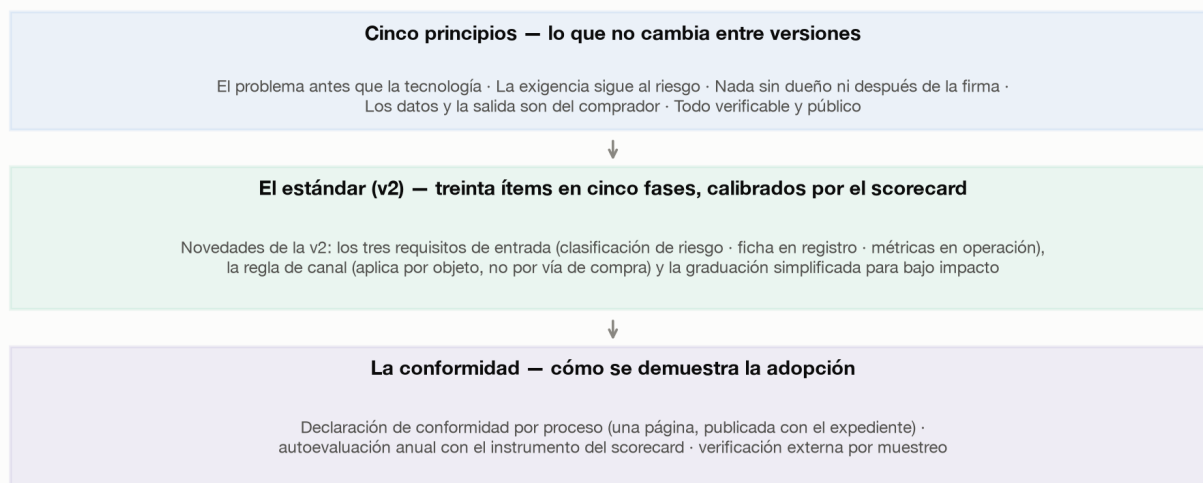
Tabla 3. El marco aplicable frente a las tres condiciones de entrada. Fuente: Enpath, con base en las normas citadas.

La lectura es la de toda esta serie: **el fundamento está completo y la exigencia de forma no existe.** La circular de la autoridad de datos incluso proporciona el test sustantivo —proporcionalidad en sentido estricto— que la clasificación de riesgo del estándar operacionaliza. Lo que falta es el acto que convierta el test en campo obligatorio del pliego.

Parte III — El estándar

7. Los cinco principios

Marco Enpath de compra responsable de IA · Tres niveles



Fuente: Enpath. El estándar detallado: informe de mayo de 2025; la evidencia de calibración: scorecard de noviembre de 2025.

Figura 1. Los tres niveles del MCR. Fuente: Enpath.

1. El problema antes que la tecnología. Ningún proceso sin problema, resultados y alternativas documentados. La fase 0 es el fundamento, y su ausencia es la causa raíz medida del resto: 21,8 % de cumplimiento medio, con dos de sus seis ítems en cero.

2. La exigencia sigue al riesgo. La matriz de impacto —no el precio, no el canal, no el proveedor— determina qué se exige. Corolario de la v2: la regla de canal.

3. Nada sin dueño, tampoco después de la firma. Todo proceso con gestor técnico (8 de 20 en la medición); toda operación con métricas, monitoreo y reevaluación con responsable. La fase 4 deja de ser tierra de nadie.

4. Los datos y la salida son del comprador. Propiedad, exportación y transición pactadas antes de la primera firma. La dependencia es una cláusula ausente, no un destino: propiedad y exportación en 35 %, plan de salida en 20 %.

5. Todo verificable y público. Requisitos con evidencia, aceptación con pruebas, ficha en el registro, declaración de conformidad publicada. **Lo que no puede mostrarse no se cuenta como cumplido.**

La regla de inmutabilidad. Cualquier revisión futura del estándar que contradiga un principio no es una versión: es otro marco. Los principios son el contrato de este documento con sus adoptantes, y la §10 define cómo se protege.

8. El estándar v2: tres condiciones de entrada, regla de canal, graduación

El detalle de los treinta ítems permanece en el informe de mayo de 2025 y en su archivo de datos. **La v2 no renumera:** añade tres condiciones de entrada, la regla de canal y la variante simplificada de doce ítems para bajo impacto.

8.1. Las tres condiciones de entrada

Figura 6 · La medición encontró cuatro ceros absolutos; el marco convirtió tres en condiciones de entrada



Fuente: scorecard Enpath de noviembre de 2025. El cuarto cero —la regla de ambición sobre madurez del proveedor— no fue elevado a condición de entrada por la versión 2 del estándar, y el documento original no explicaba por qué.

Figura 6. La medición encontró cuatro ceros absolutos; el marco convirtió tres en condiciones de entrada. Fuente: scorecard Enpath de noviembre de 2025.

Condición	Qué exige	Cero medido	Cómo se verifica (Anexo C)
CE1 · Clasificación de riesgo	El sistema clasificado en la matriz de impacto, con la clasificación en el estudio previo y publicada	0 de 20	Existencia del documento con la clasificación y su fundamento
CE2 · Ficha en el registro	Ficha pública del sistema en el registro de algoritmos de la entidad, enlazada desde el contrato	0 de 20	Existencia de la ficha con los campos mínimos
CE3 · Métricas de operación	Métricas definidas en el contrato, con umbral y periodicidad, y compromiso de publicación	0 de 17	Existencia de la cláusula con métricas, umbrales y periodicidad

Tabla 4. Las tres condiciones de entrada. Sin ellas, el proceso no se considera conforme, cualquiera sea su puntaje en los otros veintisiete ítems. Fuente: Enpath.

El diseño es deliberadamente binario y severo: **un proceso con veintisiete ítems cumplidos y sin clasificación de riesgo no es conforme.** La razón es la que la medición mostró: los ítems que quedan como «uno entre treinta» se negocian a la baja; los que son condición de entrada, no.

8.2. El cuarto cero y qué hacer con él

Esta subsección es nueva y responde a la. La medición encontró un cuarto ítem en cero absoluto: **la regla de ambición, es decir, la exigencia de contrastar la madurez de la solución y del proveedor con la complejidad del problema (0 de 20).**

No fue elevado a condición de entrada. Por qué, y qué hacer al respecto:

Por qué no es condición de entrada. Las tres condiciones se verifican por existencia de un documento: la clasificación existe o no; la ficha existe o no; la cláusula de métricas existe o no. La regla de ambición exige un **juicio sustantivo** —si la madurez ofrecida es proporcional al problema— que no puede resolverse con verificación documental binaria y que una oficina jurídica no está en posi-

ción de hacer. Convertirla en condición de entrada habría producido cumplimiento formal vacío: una frase declarativa en cada estudio previo.

Qué se hace en su lugar. Se traslada al bloque de la declaración de conformidad como **campo obligatorio de justificación**: el proceso debe declarar, en dos o tres frases, por qué la madurez de la solución elegida es proporcional a la complejidad del problema y qué alternativa de menor sofisticación se descartó. Es una obligación de argumentar, no de existir — y su valor está en que queda publicada y por tanto es criticable. El texto exacto está en el Anexo A, campo 7.

Y una regla de método: un marco que dice calibrarse con su propia medición debe contar los ceros correctamente. Contarlos mal no cambia ninguna de las tres condiciones de entrada, y sí muestra que la revisión aritmética de las cifras propias no fue suficiente.

8.3. La regla de canal

El estándar aplica **por objeto** —sistema que decide o apoya decisiones sobre personas o recursos— y no por modalidad de contratación, cuantía o canal. La razón es empírica: el peor cumplimiento medido se concentró en el canal de contratación menor. La cifra específica (14 %) no es reproducible con los datos publicados, pero la dirección del hallazgo es consistente con la lógica del canal: donde el trámite es más ligero, los requisitos técnicos desaparecen primero.

La regla cierra la vía formal de eludir el estándar bajando la cuantía. **No cierra la vía legítima**, que es la graduación.

8.4. La graduación simplificada, que ya operaba

Para sistemas de bajo impacto, el estándar aplica en su variante de **doce ítems** en lugar de treinta. El propósito es que la exigencia no empuje compras a la informalidad.

El hallazgo: **la graduación ya operaba en el terreno antes de formalizarse**. Diez de los treinta ítems del scorecard tienen denominadores inferiores a veinte —entre 6 y 17— porque no aplicaban a todos los procesos evaluados. Es decir, los evaluadores ya distinguían aplicabilidad por naturaleza del proceso. La v2 no introdujo la graduación: la escribió (§11.5).

9. La conformidad: declaración, autoevaluación y verificación externa

La pieza nueva del marco, y la que define qué significa adoptarlo.

9.1. La declaración de conformidad

Una página por proceso, publicada con el expediente. Media página de identificación —sistema, clasificación de riesgo, canal, valor— y media de cumplimiento —ítems aplicables, ítems cumplidos, evidencia enlazada, excepciones firmadas y justificación de la regla de ambición. El texto modelo completo está en el Anexo A.

Su publicación cierra el circuito con dos instrumentos de la serie: con el **registro de algoritmos** (la ficha enlaza la declaración) y con el **scorecard** (que pasa de reconstruir expedientes a auditar declaraciones — el salto de costo que haría las próximas mediciones más grandes y más baratas).

La regla de honestidad del marco: adoptar el MCR es publicar la declaración. Lo demás es citarlo.

9.2. La autoevaluación anual

La entidad se aplica el instrumento del scorecard sobre sus propios procesos del año y publica el resultado. Es el mecanismo que convierte la adopción en serie temporal propia, y el que permite a una entidad demostrar mejora sin esperar a que un tercero la mida.

9.3. La verificación externa por muestreo

El mecanismo que separa la adopción real de la retórica, y el único que contiene el riesgo de la conformidad de papel (R1). Puede ejercerlo un órgano de control, un par institucional o un tercero independiente. El protocolo está en el **Anexo D** y su regla central es que **la verificación se hace contra la evidencia enlazada en la declaración, no contra la declaración**: una declaración que afirma cumplimiento con enlaces que no abren, no documentan o no corresponden es un incumplimiento con agravante.

10. La gobernanza del marco

Un estándar sin gobernanza declarada es una opinión con formato. La del MCR tiene cuatro reglas:

1. **Principios inmutables.** Los cinco principios de la §7 no se modifican. Una revisión que contradiga uno es otro marco y debe llamarse de otro modo.
2. **Versionado explícito.** Los cambios de contenido producen versión nueva con número, fecha y justificación pública del cambio. La v2 es de enero de 2026; la revisión siguiente ampliada **no** produce v3: aclara, corrige aritmética y añade anexos, sin modificar ítems ni condiciones.
3. **Revisión anual con evidencia.** Cada revisión se apoya en la medición disponible, y **cualquier ítem que la medición muestre inútil se retira**. El estándar que no se deja corregir por los datos no puede exigir que los gobiernos sí.
4. **Registro público de erratas.** Los errores de hecho o de aritmética del marco se publican en una sección de erratas fechada, como la §2. No en notas al pie.

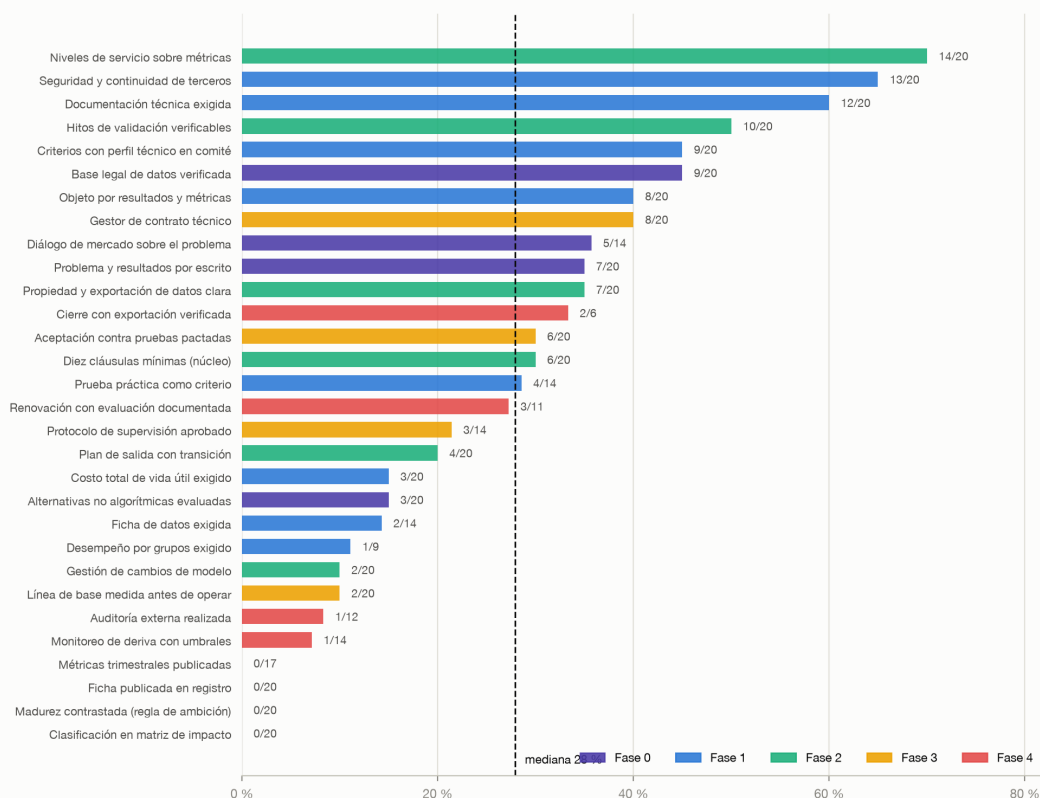
Parte IV — Evidencia

11. Análisis de datos: los treinta ítems, medidos

El scorecard no se lee aquí «como especificación»: se recomputa y se somete a prueba.

11.1. Los treinta ítems, ordenados

Figura 3 · Los treinta ítems del estándar, ordenados por cumplimiento medido
Cada etiqueta indica procesos que cumplen sobre procesos aplicables



Fuente: scorecard Enpath de compra pública de IA (noviembre de 2025), archivo publicado resultados_scorecard.csv. Los denominadores varían entre 6 y 20 por la graduación por riesgo del estándar; diez de los treinta ítems no aplican a todos los procesos.

Figura 3. Los treinta ítems del estándar, ordenados por cumplimiento medido. Cada etiqueta indica procesos que cumplen sobre procesos aplicables. Fuente: scorecard Enpath de noviembre de 2025, archivo publicado.

La figura es la fotografía completa del estado de la compra pública de IA medida en la región, ítem por ítem. Su lectura arroja tres bloques claros:

El bloque alto (por encima del 45 %): seis ítems, todos de contratación convencional. Niveles de servicio sobre métricas (70 %), seguridad y continuidad de terceros (65 %), documentación técnica exigida (60 %), hitos de validación verificables (50 %), base legal de datos verificada (45 %) y criterios con perfil técnico en comité (45 %). **Ninguno es específico de inteligencia artificial.** Son cláusulas que cualquier contrato de tecnología ya contiene.

El bloque medio (20 %-40 %): once ítems, mezcla de requisitos convencionales y de los pocos requisitos de IA que encontraron acomodo en formatos existentes: gestor de contrato técnico (40 %), objeto por resultados (40 %), diálogo de mercado (36 %), problema y resultados por escrito (35 %), propiedad y exportación (35 %), cierre con exportación verificada (33 %), aceptación contra pruebas (30 %), diez cláusulas mínimas (30 %), prueba práctica como criterio (29 %), renovación con evaluación (27 %), protocolo de supervisión (21 %).

El bloque bajo (por debajod del 20 %): trece ítems, casi todos específicos de sistemas algorítmicos. Plan de salida (20 %), alternativas no algorítmicas (15 %), costo total de vida útil (15 %), ficha de datos exigida (14 %), desempeño por grupos (11 %), línea de base medida (10 %), gestión de cambios de modelo (10 %), auditoría externa (8 %), monitoreo de deriva (7 %) y **los cuatro ceros**.

11.2. El descenso por fase

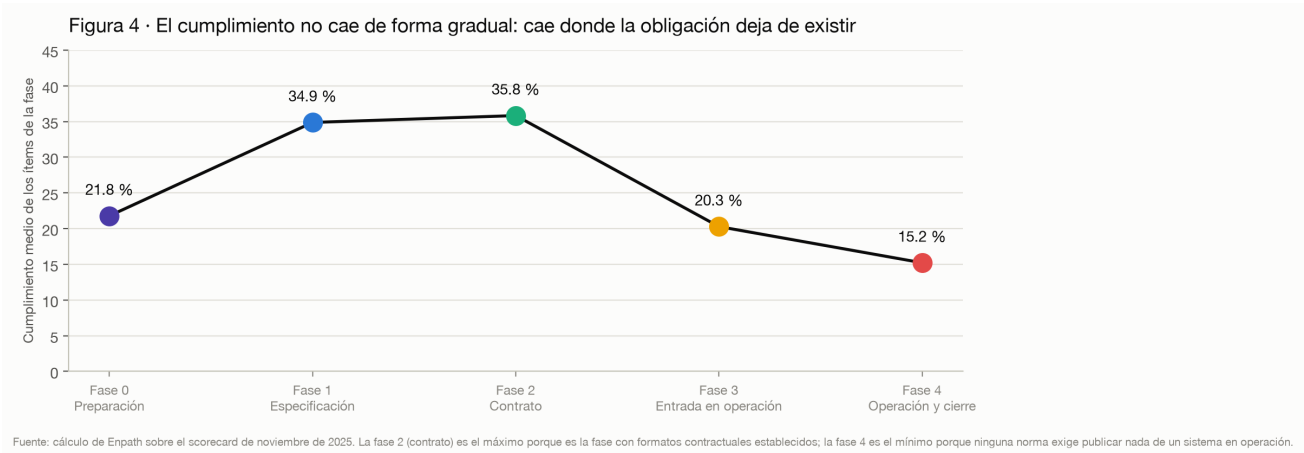


Figura 4. El cumplimiento no cae de forma gradual: cae donde la obligación deja de existir. Fuente: cálculo de Enpath sobre el scorecard de noviembre de 2025.

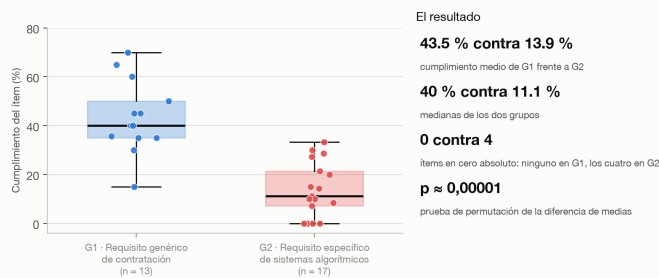
Fase	Íte- ms	Cumplimiento me- dio	Lectura
Fase 0 · Preparación	6	21,8 %	Nadie tiene formulario para clasificar riesgo ni para documentar alter- nativas
Fase 1 · Especificación	8	34,9 %	Los pliegos tienen formatos; los requisitos de IA que caben en ellos se cumplen
Fase 2 · Contrato	6	35,8 % — el máxi- mo	La fase con formatos contractuales más establecidos
Fase 3 · Entrada en opera- ción	5	20,3 %	La transición entre el contrato y la operación no tiene dueño
Fase 4 · Operación y cierre	5	15,2 % — el mini- mo	Ninguna norma exige publicar nada de un sistema en funcionamiento

Tabla 5. Cumplimiento por fase. Fuente: cálculo de Enpath sobre el scorecard de noviembre de 2025.

La forma de la curva es el argumento. No es un descenso monótono a lo largo del ciclo —lo que su-
geriría fatiga o pérdida de atención—: es un máximo en la fase con más formatos y dos mínimos en
las fases sin ninguno. La preparación y la operación son las dos fases donde el régimen de
contratación no prescribe documentos, y son exactamente las dos peores. **El cumplimiento sigue
la disponibilidad de formularios, no el orden temporal.**

11.3. La hipótesis de los formularios, puesta a prueba

Figura 5 · La hipótesis de los formularios, puesta a prueba: los estándares se cumplen donde ya hay formato



Fuente: cálculo de Enpath sobre el scorecard de noviembre de 2025, con clasificación declarada de los treinta ítems en G1 (requisito que existe en los formatos habituales de contratación tecnológica) y G2 (requisito específico de sistemas algorítmicos). La clasificación completa se publica en el Anexo B para que pueda disputarse.

Figura 5. La hipótesis de los formularios, puesta a prueba: los estándares se cumplen donde ya hay formato.
Fuente: cálculo de Enpath sobre el scorecard de noviembre de 2025, con la clasificación declarada del Anexo B.

El diseño de la prueba. Los treinta ítems se clasificaron en dos grupos según un criterio único y declarado de antemano: ¿el requisito existe ya en los formatos habituales de contratación tecnológica?

- **G1 (13 ítems)** — requisitos genéricos de contratación tecnológica: problema y resultados por escrito, diálogo de mercado, base legal de datos, objeto por resultados, documentación técnica, criterios con perfil técnico, seguridad y continuidad, costo total de vida útil, niveles de servicio, hitos de validación, propiedad y exportación, gestor de contrato, aceptación contra pruebas.
- **G2 (17 ítems)** — requisitos específicos de sistemas algorítmicos: clasificación en matriz, regla de ambición, alternativas no algorítmicas, ficha de datos, desempeño por grupos, prueba práctica, diez cláusulas mínimas, gestión de cambios de modelo, plan de salida, protocolo de supervisión, línea de base antes de operar, ficha en registro, métricas trimestrales, auditoría externa, monitoreo de deriva, renovación con evaluación, cierre con exportación verificada.

La clasificación completa, con el criterio aplicado a cada ítem, está en el **Anexo B**, publicada precisamente para que pueda disputarse y el cálculo repetirse con otra codificación.

El resultado.

Grupo	n	Media	Mediana	Ítems en cero
G1 · Requisito genérico	13	43,5 %	40,0 %	0
G2 · Específico de sistemas algorítmicos	17	13,9 %	11,1 %	4
Diferencia de medias	—	29,6 pp	—	—

Tabla 6. La prueba de la hipótesis de los formularios. Prueba de permutación de la diferencia de medias, 200.000 remuestreos, una cola: $p \approx 0,00001$. Fuente: cálculo de Enpath.

La interpretación, con sus límites. El resultado es consistente con la hipótesis del marco: los requisitos se cumplen donde ya existe un formulario donde escribirlos. Pero conviene declarar tres advertencias antes de usarlo como prueba:

1. **La clasificación es propia y podría estar contaminada por el resultado.** Se hizo con un criterio declarado y sin mirar los porcentajes, pero el equipo conocía la medición. La mitigación es publicar la clasificación (Anexo B) para que un tercero la recodifique.
2. **Correlación no es mecanismo.** Es posible que los requisitos específicos de IA se cumplan menos no porque falte el formulario, sino porque son intrínsecamente más difíciles o más costosos. La observación que favorece la hipótesis del formulario: varios ítems de G2 son **más baratos** que ítems de G1 bien cumplidos —publicar una ficha cuesta menos que negociar niveles de servicio— y aun así cumplen mucho menos.
3. **Veinte procesos no son una muestra probabilística.** El resultado es una regularidad en la muestra examinada, no una estimación poblacional.

La predicción falsable que el mecanismo implica: los ítems que entren a los formatos de las agencias centrales subirán del cero a la mayoría en un ciclo de contratación; los que queden como recomendación seguirán donde están. Es contrastable en la segunda edición del scorecard.

11.4. Los cuatro ceros y su naturaleza común

Los cuatro ítems en cero absoluto comparten un rasgo que los tres primeros bloques no capturan: **ninguno tiene un lugar en ningún documento que la administración ya produzca.** La clasificación de riesgo no cabe en el estudio previo tradicional; la regla de ambición no cabe en la evaluación de ofertas; la ficha en el registro presupone un registro que no existe; las métricas trimestrales presuponen un canal de publicación que nadie definió.

Es la razón por la que las tres condiciones de entrada del estándar están redactadas como **existencia de documento** y no como calidad de contenido: en la fase en que un requisito vale cero, exigir calidad es prematuro. Primero tiene que existir el campo.

11.5. La graduación ya operaba

Denominador	Ítems	Interpretación
20 (todos los procesos)	20	Requisitos de aplicación universal
17	1	Procesos que llegaron a fase de operación
14	3	Aplicabilidad por naturaleza del sistema
12	1	Procesos con antigüedad suficiente para auditoría
11	1	Procesos con decisión de renovación
9	1	Sistemas que deciden sobre personas (desempeño por grupos)
6	1	Procesos terminados (cierre con exportación)

Tabla 7. Denominadores de los treinta ítems. Diez ítems tienen aplicabilidad restringida. Fuente: scorecard Enpath de noviembre de 2025.

El hallazgo: un tercio de los ítems del estándar ya se aplicaba de forma graduada en el terreno, antes de que la v2 escribiera la graduación. Es un argumento a favor de su viabilidad, y una advertencia metodológica para leer la Figura 3: **los porcentajes no son comparables entre ítems con denominadores distintos**, y los ítems de la fase 4 tienen los denominadores más pequeños, lo que hace sus cifras más frágiles.

12. Comparación internacional y estudios de caso



Figura 7. El MCR frente a sus pares: dos casillas propias y una casilla vacía. Fuente: Enpath, con base en los instrumentos citados.

Criterio	MCR (ene. 2026)	Cláusulas modelo UE (mar. 2025)	EE. UU. M-25-22 (abr. 2025)	Chile · Bases Tipo (ene. 2023)	WEF (jun. 2020)
Cubre la fase previa (0)	Sí	Parcial	Sí	Parcial	Sí
Cláusulas contractuales	Sí	Sí	Sí	Sí	Parcial
Cubre operación y cierre	Sí	Parcial	Sí	No	Parcial
Conformidad publicada	Sí	No	No	No	No
Instrumento de medición propio	Sí	No	Parcial	Parcial	No
Adoptado por agencia central	No	Sí	Sí	Sí	No

Tabla 8. El MCR frente a cuatro pares. Fuente: Enpath, con las fuentes primarias citadas.

12.1. Frente a las cláusulas modelo europeas

Las cláusulas contractuales modelo para la compra pública de IA, en su versión actualizada del 5 de marzo de 2025, ofrecen versión completa para alto riesgo, versión ligera para no alto riesgo y comentario de uso, siguiendo en lo sustancial las obligaciones del capítulo III del Reglamento (UE) 2024/1689².

La diferencia se mantiene: el MCR cubre el ciclo completo —fases 0, 3 y 4— y las cláusulas cubren el contrato. **Y el MCR reconoce su deuda:** la graduación por riesgo del estándar europeo es el mo-

delo de la propia graduación del MCR, y la versión ligera europea es el antecedente directo de la variante de doce ítems.

La recomendación práctica que no cambia: una entidad de la región que adopte las cláusulas modelo europeas y les añada las tres condiciones de entrada y la declaración de conformidad tiene, en dos movimientos, el estándar completo — con un texto redactado por juristas de compra pública y que los proveedores globales reconocen.

12.2. La asimetría documental, con dieciséis meses más

El argumento —la industria global produce por obligación europea la documentación que la región no exige— se sostiene con dos precisiones de calendario. Las obligaciones de propósito general europeas son aplicables desde **agosto de 2025**. El régimen de alto riesgo, en cambio, se desplazó: **Anexo III al 2 de diciembre de 2027 y Anexo I al 2 de agosto de 2028**, tras el Digital Omnibus adoptado por el Parlamento Europeo el 16 de junio de 2026 y por el Consejo el 29 de junio⁴.

Traducción práctica actualizada: exigir el estándar sigue siendo barato, porque el mercado global lo construirá de todos modos, y **la ventana para exigirlo sin fricción con los proveedores es ahora más larga, no más corta**. El riesgo del aplazamiento no es técnico sino de agenda: que se lea como permiso para esperar (§14, R7).

12.3. Frente al régimen federal estadounidense

El memorando de adquisición de IA fue emitido el **3 de abril de 2025** y derogó y reemplazó al de 2024. Se organiza en tres ejes: evitar la dependencia de proveedor y promover la interoperabilidad; proteger el gasto público mediante seguimiento del desempeño de la IA y gestión de riesgos; y promover la adquisición efectiva con equipos interfuncionales. **Cubre el ciclo completo de adquisición**, desde la identificación de requisitos y el estudio de mercado hasta la selección, la administración del contrato y su cierre; exige validación de desempeño y pruebas previas al fallo para IA de alto impacto; y protege derechos sobre datos y propiedad intelectual, con plazo de cumplimiento pleno de 180 días³.

La corrección al original. El documento de enero afirmaba que el MCR se distingue del régimen europeo por cubrir el ciclo completo y del estadounidense por añadir la conformidad publicada. Lo segundo es correcto; lo primero necesita el matiz de que **el régimen estadounidense también cubre el ciclo completo, incluido el cierre**. La distinción real del MCR frente a Estados Unidos es una sola: **la práctica federal documenta hacia adentro; el MCR, hacia el público**.

La lección transferible, y es la más útil de las cuatro comparaciones: el memorando exige **pruebas previas al fallo para IA de alto impacto**. El ítem equivalente del estándar —prueba práctica como criterio de evaluación— cumple en la región el 29 % (4 de 14). Es un requisito con precedente federal, formato claro y costo bajo, y es un candidato natural a condición de entrada en una futura v3.

12.4. Caso Chile: la agencia central que ya adoptó un estándar de compra de IA

Es el caso más relevante de toda la comparación para un lector latinoamericano, y el que con más frecuencia se subestima.

Qué existe. ChileCompra —la Dirección de Compras y Contratación Pública— publicó **Bases Tipo para la licitación de proyectos de algoritmos e inteligencia artificial con requisitos éticos**, dispo-

nibles desde **enero de 2023**. El proceso fue abierto: consulta pública del 5 de octubre al 18 de noviembre de 2022 y respuestas a los comentarios publicadas el 2 de diciembre de 2022. Las bases incorporan hitos de **transparencia, privacidad, no discriminación y explicabilidad**, y exigen —según la naturaleza del contrato— el uso de modelos con **métricas estadísticas de equidad**, medidas adicionales de protección de datos y **análisis de sesgo**¹.

De dónde salió. Del proyecto «Algoritmos Éticos, Responsables y Transparentes», ejecutado por el **GobLab de la Universidad Adolfo Ibáñez** con apoyo de **BID Lab** desde diciembre de 2020, con ChileCompra, el Ministerio de Ciencia, la División de Gobierno Digital de la Secretaría General de la Presidencia, el Consejo para la Transparencia y varias entidades prestadoras como socios. Entre 2023 y 2025 el proyecto publicó como bienes públicos tres instrumentos: la **Evaluación de Impacto Algorítmico**, la herramienta de **medición de sesgo y equidad** y la **Ficha de Transparencia Algorítmica**. A julio de 2024 acumulaba cuatro reconocimientos internacionales, y en diciembre de 2025 ChileCompra aplicó la evaluación de impacto a su propio sistema de detección de anomalías en procesos de compra¹.

Qué cubre y qué no, en términos del MCR.

Elemento del MCR	¿Presente en el caso chileno?
Requisitos éticos exigibles en el pliego	Sí, estandarizados y en bases tipo de la agencia central
Clasificación de riesgo (CE1)	Parcial: existe la evaluación de impacto algorítmico como instrumento
Ficha pública del sistema (CE2)	Parcial: existe la Ficha de Transparencia Algorítmica como instrumento
Métricas de operación publicadas (CE3)	No
Declaración de conformidad publicada por proceso	No
Instrumento de medición con línea de base	Parcial: no se identificó una medición de cumplimiento equivalente al score-card
Adopción por agencia central	Sí — desde enero de 2023

Tabla 9. El caso chileno frente a los elementos del MCR. Fuente: Enpath, con las fuentes citadas.

Las cuatro lecciones.

- 1. La ruta 2 es viable y ya se recorrió.** La apuesta del MCR para 2026 —que una agencia central adopte un estándar de compra de IA— describe algo que un vecino hizo en 2023. Eso cambia el argumento con el que se debe convencer a la agencia colombiana: no «es posible», sino «se hizo, con consulta pública, y ha sido premiado internacionalmente».
- 2. La combinación institucional que funcionó fue academia + banca de desarrollo + agencia central.** No una ley, ni una política nacional, ni una estrategia. Es un modelo de producción de estándares replicable sin esperar al legislador — y es, notablemente, el modelo en el que Enpath ocupa la posición de la academia sin haber conseguido todavía la de la agencia central.
- 3. La consulta pública fue parte del método.** Someter las bases a comentario abierto y publicar las respuestas es lo que convierte un instrumento técnico en un estándar de mercado, porque los proveedores participan antes de tener que cumplir. **El MCR no ha hecho consulta pública**, y esta edición lo registra como carencia (§16.5).

4. **La conformidad publicada sigue siendo el hueco, incluso en el mejor caso regional.** Las Bases Tipo chilenas no incluyen declaración de conformidad por proceso ni publicación de métricas de operación. La contribución específica del MCR se estrecha y se aclara: **no es el estándar de compra de IA de la región; es el que añade la demostración pública de adopción y el instrumento para medirla.**

12.5. Frente al instrumento del Foro Económico Mundial

Publicado el **11 de junio de 2020**, no es solo un marco de principios: incluye guías de adquisición, un cuaderno de trabajo para funcionarios de política y compras y lecciones de implementación. Se pilotó con departamentos y agencias de Bahreín y el Reino Unido y con la empresa de electricidad y agua de Dubái, **con retroalimentación de pilotos en Colombia** y otros gobiernos socios, y está concebido como documento vivo⁵.

Un dato que importa para la historia nacional: Colombia participó como piloto de retroalimentación en 2020. Es decir, la agenda colombiana de compra responsable de IA no empezó en 2024 con esta serie: empezó seis años antes, en un piloto internacional cuyo rastro institucional no es visible en la medición de 2025. **Un país que pilotó guías internacionales de compra de IA en 2020 tenía, cinco años después, cero de veinte procesos con clasificación de riesgo.** Es la ilustración más nítida de la brecha instrumentos-hábitos que toda esta serie documenta.

13. Análisis económico y de actores

13.1. El costo de adoptar contra el costo de no adoptar

Ruta	Esfuerzo	Plazo	Alcance	Costo directo
Ruta 1 · Entidad	Una resolución, una plantilla, capacitación de un equipo	Semanas	Una entidad	Marginal: días-persona de la oficina jurídica y de planeación
Ruta 2 · Agencia central	Documentos tipo, consulta pública, ajuste de plataformas	Meses	La mediana nacional	Bajo: es trabajo normativo ordinario de la agencia
Ruta 3 · Regional	Acuerdo entre agencias, condiciones de banca multilateral	Años	El mercado regional	Bajo por país, alto en coordinación

Tabla 10. Costo, plazo y alcance de las tres rutas. Estimación cualitativa de Enpath. Fuente: Enpath.

Y el costo de no adoptar, que sí está medido. La mediana del 27,9 % no es una cifra abstracta: es un conjunto de contratos vigentes, con duraciones de años, en los que ningún sistema tiene clasificación de riesgo publicada, ninguno tiene ficha pública, ninguno publica métricas, uno de doce ha sido auditado y tres de once renovaciones se decidieron con evaluación. Cada mes de demora en las rutas 1 y 2 añade contratos a ese inventario, y **la ventana de corrección de un contrato es su firma.**

13.2. Actores

Actor	Qué gana con el MCR	Qué pierde	Comportamiento previsible
Entidad compradora	Criterio para exigir y defenderse; menos dependencia	Velocidad aparente en el proceso	Favorable si la adopción es una plantilla, adversa si es un proyecto
Oficina jurídica	Texto adoptado en lugar de negociación caso a caso	—	Aliado natural
Área de tecnología	Requisitos técnicos con respaldo institucional	Discrecionalidad	Favorable
Agencia central de compras	Instrumento de política de alto alcance y bajo costo; precedente regional verificado	Costo de consulta pública y de mantenimiento	El actor decisivo, y el que no ha actuado
Proveedores serios	Previsibilidad; expediente preparado una vez	La venta por opacidad	Favorables a mediano plazo
Proveedores que compiten por opacidad	—	Ventaja competitiva	Adversos
Órganos de control	Auditoría contra declaración publicada: el control más barato disponible	—	La palanca más subestimada
Banca multilateral	Condición de desembolso verificable desde el portal	—	Favorable; el caso chileno demuestra que ya financia esta agenda
Ciudadanía y prensa	La declaración como puerta de entrada al expediente	—	Uso probable si existe el documento

Tabla 11. Análisis de actores. Fuente: Enpath.

La conclusión del mapa es la misma que la Figura 7 exhibe en forma de casilla vacía: **el actor decisivo es la agencia central de compras y es el único que no ha actuado**. Todos los demás tienen incentivos alineados o costos marginales. La intervención con mejor relación costo-impacto que no depende de esa agencia es la de los órganos de control: auditar contra declaraciones publicadas no requiere norma nueva, y su mera anunciación produciría declaraciones.

Parte V — Prospectiva y decisión

14. Riesgos

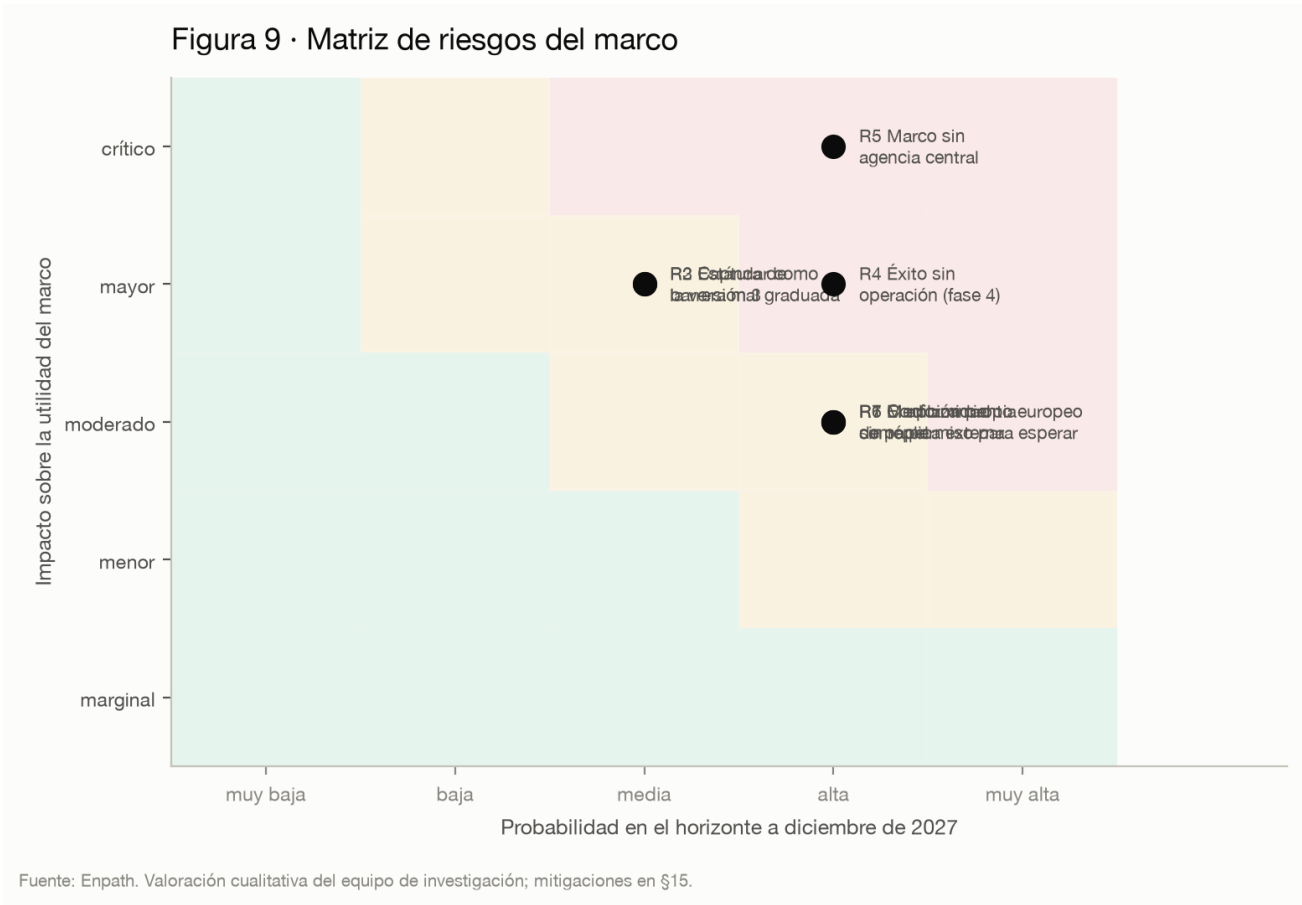


Figura 9. Matriz de riesgos del marco. Valoración cualitativa del equipo de investigación. Fuente: Enpath.

Id	Riesgo	Pro-babili-dad	Im-pacto	Mitigación de diseño
R1	La conformidad de papel: declaraciones publicadas con evidencia débil	Alta	Mo-dera-do	La verificación por muestreo (Anexo D), cuya regla central es auditar la evidencia enlazada y no la declaración. El marco asume que habrá simulación; su diseño la hace medible
R2	El estándar como barrera mal graduada: si la variante simplificada no se usa, la exigencia plena empuja compras a la sombra	Media	Ma-yor	La regla de canal cierra la puerta formal; la graduación mantiene abierta la legítima — y la \$11.5 muestra que ya operaba en el terreno
R3	La captura de la v3: con adopción real vendrá presión por diluir ítems incómodos (auditoría, salida, métricas)	Media	Ma-yor	Principios inmutables y cambios versionados con justificación pública (§10); la vigilancia de los adoptantes es el complemento
R4	El éxito sin operación: documentos tipo que suben las fases 0–2 y dejan la 4 en el desierto medido del 15,2 %	Alta	Ma-yor	La condición de entrada CE3 (métricas de operación) existe para eso; su aplicación es la prueba real del marco
R5	El marco sin agencia central: adopción de entidad dispersa, sin efecto sobre la mediana nacional	Alta	Críti-co	Es la casilla vacía de la Figura 7 y el riesgo principal. Mitigación disponible: el precedente chileno verificado (§12.4) y la consulta pública pendiente (§16.5)
R6	La medición propia sin réplica externa: Enpath diseña el estándar, lo mide y lo audita	Alta	Mo-dera-do	Metodología e instrumento publicados; datos publicados y recomputados; el Anexo B publica incluso la codificación del análisis para que se dispute
R7	El aplazamiento europeo como permiso para esperar: el desplazamiento del régimen de alto riesgo a diciembre de 2027 invocado para postergar exigencias contractuales	Alta	Mo-dera-do	El argumento de la §12.2: el contenido del estándar no cambió y la ventana de exigirlo sin fricción es ahora más larga. La predicción P4 (§17) está formulada para detectar si el riesgo se materializa

Tabla 12. Registro de riesgos. Fuente: Enpath.

Dos comentarios. R5 es el riesgo crítico y define el estado real del marco a julio de 2026: un estándar bien construido, calibrado con datos propios, con instrumento de medición y sin la adopción que lo haría relevante. Y R6 merece una nota de honestidad: este marco no resuelve el problema de que Enpath ocupe simultáneamente los papeles de autor, medidor y auditor. Lo único que puede hacer un autor en esa posición es **entregar todo lo necesario para ser contradicho** — los datos, la codificación, el protocolo y la lista de sus propios errores. Eso es lo que la §2 y los anexos hacen.

15. Oportunidades

O1. La adopción es un acto administrativo, no un proyecto. La ruta de entidad cuesta una resolución y una plantilla; el retorno —salir de la banda medida del 28 %— es el más barato de toda la agenda. El Anexo A entrega el texto.

O2. La declaración de conformidad como señal de mercado. Los proveedores serios la usarán como argumento comercial; los compradores, como filtro previo. El estándar crea el equilibrio que castiga la promesa y premia la evidencia.

O3. El scorecard 2026–2027 como árbitro barato. Con declaraciones publicadas, la próxima medición podrá triplicar la muestra a una fracción del costo — y atribuir el cambio a las rutas adoptadas.

O4. El precedente chileno como argumento, no como competencia. La existencia de Bases Tipo chilenas desde 2023 retira el argumento de inviabilidad y ofrece un modelo institucional replicable: academia + banca de desarrollo + agencia central. **Es la mejor noticia de este documento para la agenda regional, aunque sea la peor para la pretensión de novedad del marco.**

O5. La primera agencia central que adopte se lleva el estándar regional. La dinámica documentada de fichas y registros aplica: quien publique primero los documentos tipo fija el formato que los demás adaptarán. Chile lo hizo con las bases; el formato de la conformidad publicada sigue vacante.

O6. Las pruebas previas al fallo como candidata a v3. Con precedente federal estadounidense explícito, formato claro y costo bajo, y un cumplimiento regional del 29 %, es el ítem con mejor relación entre respaldo internacional y margen de mejora.

O7. La ventana contractual más larga. El aplazamiento europeo da dieciséis meses adicionales para exigir por contrato la documentación que el mercado global producirá de todos modos, sin fricción con proveedores que aún no están obligados.

16. Recomendaciones y las tres rutas de adopción

Tres rutas de adopción — acumulables, no excluyentes

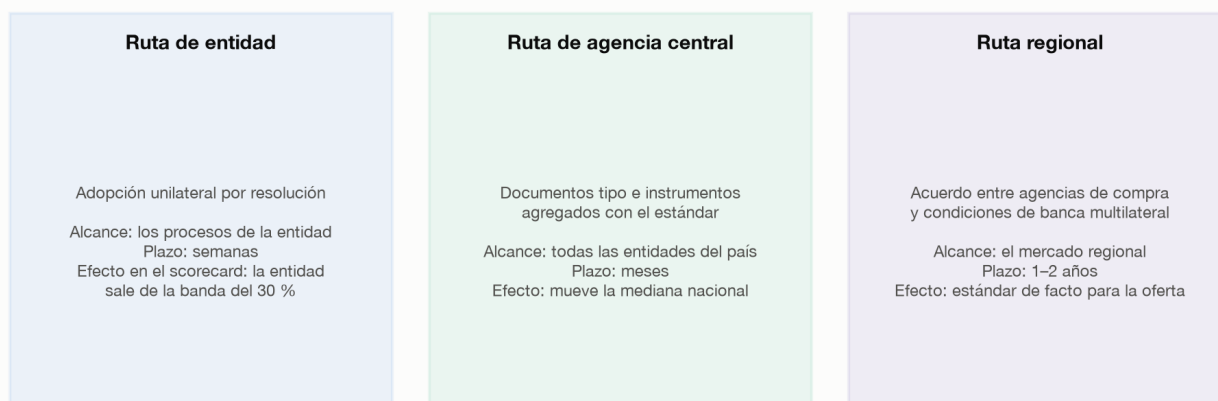


Figura 2. Las tres rutas de adopción. Fuente: Enpath.

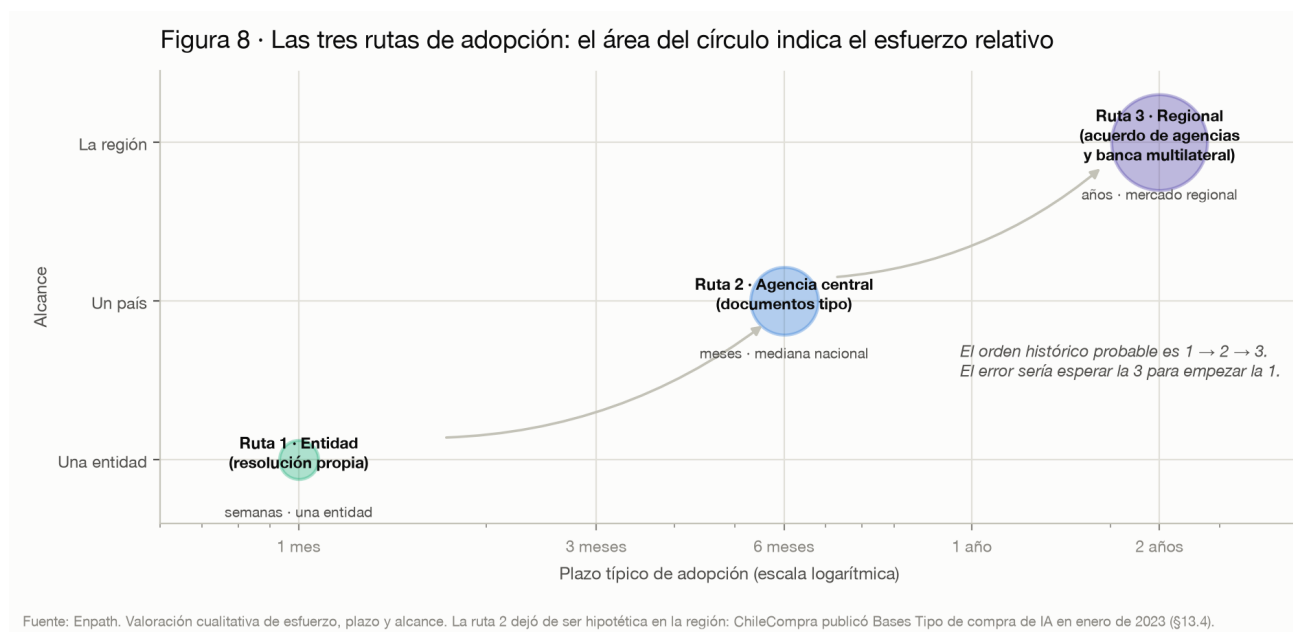


Figura 8. Las tres rutas de adopción: el área del círculo indica el esfuerzo relativo. Fuente: Enpath.

16.1. Para las entidades

1. **Adopten por resolución este trimestre: estándar v2, declaración de conformidad en cada proceso, autoevaluación anual.** El texto de la declaración está en el Anexo A y los criterios de verificación de las tres condiciones de entrada en el Anexo C.
2. **Empiencen por las tres condiciones de entrada, no por los treinta ítems.** Si la adopción completa parece mucha negociación interna, tres condiciones no lo son — y son exactamente donde la medición encontró los ceros.
3. **Publiquen la primera declaración aunque el proceso no sea conforme.** Una declaración que dice «no conforme: falta CE3» es información pública valiosa y sitúa a la entidad por delante de cualquiera que no publique nada.

16.2. Para las agencias centrales de compra

Documentos tipo con las tres condiciones de entrada como mínimo. Si treinta ítems son mucha negociación, tres condiciones no lo son, y mueven los ceros medidos. Y con dos argumentos adicionales: el precedente chileno de 2023 (§12.4) demuestra viabilidad regional con consulta pública, y el memorando federal estadounidense de 2025 demuestra que las pruebas previas al fallo son exigibles en régimen real.

16.3. Para la banca multilateral

La declaración de conformidad como condición de desembolso en operaciones con componentes de IA — la ruta regional más corta disponible. Con la observación de que esta banca ya financió la producción del instrumento chileno: el papel no es nuevo, solo su aplicación como condición.

16.4. Para los demás actores

Proveedores: preparen el expediente estándar una vez —documentación, ficha de datos, plan de salida, evidencia de desempeño—; el mercado que viene lo pedirá en cada proceso. **Órganos de**

control: auditen contra la declaración; el muestreo de declaraciones publicadas es la auditoría de compra de IA de menor costo jamás disponible, y el Anexo D es el protocolo. **Todos:** midan con el instrumento común; el scorecard es replicable y sus datos comparables, y la serie temporal regional se construye entre muchos o no se construye.

16.5. Para Enpath: las tres tareas pendientes

La honestidad del marco exige que sus autores también reciban recomendaciones.

1. **Someter el MCR a consulta pública**, como ChileCompra hizo con sus bases en 2022. Un estándar que no ha sido comentado por los proveedores que deberán cumplirlo no ha sido probado donde importa.
2. **Publicar el corte por canal de contratación del scorecard**. Una de las tres calibraciones del estándar se apoya en una cifra que los datos publicados no permiten reproducir.
3. **Incluir procesos chilenos en la segunda edición del scorecard**. El contraste entre la mediana colombiana del 27,9 % y el cumplimiento de procesos regidos por Bases Tipo desde 2023 sería el dato más útil de toda la agenda regional de compra — y la prueba más directa de la hipótesis de los formularios en un régimen real.

17. Perspectivas y predicciones falsables

Con corte a julio de 2026: la apuesta de noviembre de 2025 sigue en pie y sin tomar; el índice de gobernanza de marzo de 2026 midió la adopción de entidad en Colombia y encontró ocho de ciento ocho obligaciones demostradas; y los estándares de transparencia de mayo de 2026 completaron el flanco de la fase 4.

#	Predicción	Cómo se falsea
P1	Los ítems que entren a los formatos de las agencias centrales subirán del cero a la mayoría en un ciclo de contratación; los que queden como recomendación seguirán donde están	Si un ítem sube sustancialmente sin entrar a ningún formato, la hipótesis de los formularios es insuficiente
P2	Ninguna agencia central colombiana adoptará documentos tipo con las tres condiciones de entrada antes de diciembre de 2026	Si los adopta, la predicción falla y el caso se documentará
P3	La segunda edición del scorecard encontrará la fase 4 todavía por debajo del 25 %, incluso si las fases 0–2 mejoran	Si la fase 4 supera el 25 %, el diagnóstico del «éxito sin operación» estaba sobreestimado
P4	El aplazamiento europeo al 2 de diciembre de 2027 será invocado en la región como razón para postergar exigencias contractuales de documentación	Si no aparece esa invocación en ningún expediente o comunicación pública, el riesgo R7 estaba sobreestimado
P5	Si una entidad publica declaraciones de conformidad durante un año, su cumplimiento medido subirá al menos veinte puntos respecto de su línea de base	Si publica y no sube, la declaración es un formalismo y el marco debe rediseñarla

Tabla 13. Predicciones falsables. Fuente: Enpath.

La gobernanza del propio marco queda comprometida: revisión anual con la evidencia disponible, cambios versionados, retiro de cualquier ítem que la medición muestre inútil, y registro público de erratas. **El estándar que no se deja corregir por los datos no puede exigir que los gobiernos sí.**

18. Metodología, calidad de la evidencia y limitaciones

18.1. Objetivo y preguntas

Objetivo. Consolidar la agenda de compra pública de IA de la serie en un estándar versionado, con principios estables, contenido calibrado por medición y mecanismo de demostración de adopción.

Preguntas. (i) ¿Qué contenido del estándar resistió la medición y qué exigió calibración? (ii) ¿Por qué se cumplen algunos requisitos y no otros? (iii) ¿Cómo demuestra una entidad que adoptó el estándar? (iv) ¿Qué rutas de adopción existen y con qué costo? (v) ¿Cómo se compara con los estándares internacionales vigentes?

18.2. Diseño

Marco consolidado con corte al 31 de enero de 2026, con verificación de contexto al 24 de julio de 2026. Integra sin reformular el contenido publicado y calibrado de la serie (agosto de 2024; mayo de 2025; noviembre de 2025). Las novedades de la v2 derivan de los hallazgos del scorecard. **Lo propio de este documento** es el recómputo íntegro del archivo publicado del scorecard, la prueba estadística de la hipótesis de los formularios (§11.3), la verificación de los cuatro instrumentos internacionales de comparación contra sus fuentes oficiales, y los cuatro anexos operativos.

18.3. Calidad de la evidencia

Nivel	Definición	Aplicado a
A	Medición documental primaria propia con datos publicados	Los treinta ítems del scorecard y todas las cifras derivadas (§1.1, §11)
A	Documento oficial primario	Cláusulas modelo europeas; memorando federal M-25-22; Bases Tipo de Chile-Compra; instrumento del Foro Económico Mundial; Digital Omnibus; circular de la autoridad de datos; CONPES 4144
B	Cálculo propio sobre datos de nivel A con codificación declarada	La clasificación G1/G2 y la prueba de permutación (§11.3, Anexo B)
C	Estimación cualitativa	Costos, plazos y alcances de las tres rutas (Tabla 11)
D	Valoración cualitativa estructurada	Análisis de actores; matriz de riesgos; Figura 7 (grados de cobertura de los cinco estándares)
No reproducible desde el archivo	Dato de la medición sin respaldo en los archivos publicados	El corte por canal de contratación (14 %)
Diseño	Formulación propia, no evidencia	Los cinco principios, las tres condiciones de entrada, la regla de canal, la graduación, las tres piezas de conformidad, los anexos A a D

Tabla 14. Escala de calidad de la evidencia. Fuente: Enpath.

18.4. Qué contaría como error de este informe

1. Si una recodificación razonable de los treinta ítems en G1/G2 elimina la diferencia de la §11.3, la hipótesis de los formularios no está sostenida por estos datos.

2. Si una entidad adopta el marco, publica declaraciones durante un año y su cumplimiento medido no mejora, la declaración de conformidad no produce el efecto que se le atribuye.
3. Si la variante simplificada de doce ítems resulta inaplicable o insuficiente en un caso real, la graduación está mal calibrada.
4. Si un proceso conforme según el MCR produce un sistema con fallas graves de gobernanza, el estándar tiene un vacío de contenido que debe identificarse y publicarse.
5. Si aparece un estándar de compra de IA con conformidad publicada anterior al MCR, la reivindicación de contribución específica de este marco es falsa. **El caso chileno obliga a decir esto con humildad: una primera pasada de verificación ya subestimó un instrumento vecino, y no hay razón para suponer que no existan otros.**

18.5. Limitaciones

La declaración de conformidad y la verificación por muestreo no han sido probadas en campo. Su primer ciclo de uso las calibrará. Es la limitación principal del marco.

El marco hereda las limitaciones de sus componentes, en particular la muestra no representativa del scorecard como base de calibración: veinte procesos seleccionados por variedad de sector y cuantía, no por muestreo probabilístico. **Suficiente para diseñar, insuficiente para estimar.**

Los porcentajes por ítem no son comparables entre sí cuando los denominadores difieren, y diez de los treinta ítems tienen denominadores menores que veinte (§11.5). Los ítems de la fase 4 tienen los denominadores más pequeños —entre 6 y 17—, lo que hace sus cifras las más frágiles del conjunto, precisamente donde el hallazgo es más severo.

La clasificación G1/G2 es propia y fue hecha por un equipo que conocía la medición (§11.3, advertencia 1). Se publica en el Anexo B para permitir la recodificación independiente.

El corte por canal de contratación no es reproducible con los archivos publicados.

La comparación internacional cubre cuatro instrumentos, elegidos por relevancia y disponibilidad documental. La facilidad con que el caso chileno se subestima demuestra que la revisión de instrumentos regionales fue insuficiente, y que puede seguir siéndolo.

Ninguna entidad ha adoptado el MCR al corte de este documento. Todo lo que este documento afirma sobre su funcionamiento es diseño, no observación.

18.6. Sesgos declarados

1. **Sesgo de autor-medidor-auditor.** Enpath escribió el estándar, produjo la medición que lo calibra y propone el mecanismo que lo auditaría. Es la limitación estructural del proyecto. Mitigación disponible: publicar datos, codificación, protocolo y erratas — hecho en esta edición.
2. **Sesgo de confirmación en la codificación.** La clasificación G1/G2 se hizo con criterio declarado pero por un equipo que conocía los resultados (§11.3).
3. **Sesgo de novedad.** Un marco que se presenta como consolidación regional tiene incentivo a no buscar instrumentos vecinos equivalentes. El caso chileno es la prueba de que el incentivo operó, y la §18.4 punto 5 es el compromiso de corregir si aparecen otros.
4. **Sesgo de disponibilidad en la comparación internacional.** Los instrumentos comparados son los mejor documentados en línea, con predominio de fuentes en inglés.

5. **Sesgo de agenda.** La serie sostiene que la obligación con formato es la variable determinante. La §11.3 la somete a prueba, y la §18.4 punto 1 declara cómo podría refutarse.

19. Conclusión

Un estándar merece llamarse estándar cuando tiene contenido estable, evidencia de calibración y forma de demostrar su adopción. El Marco Enpath de compra responsable de IA tiene las tres: cinco principios que no cambian, treinta ítems calibrados por una medición propia de veinte procesos reales, y una declaración de conformidad de una página que convierte la adopción en un hecho público verificable.

Y hay prueba de que la calibración fue correcta. Los estándares de compra funcionan por formularios y no por convicciones; los datos lo confirman con una nitidez inusual. Los requisitos que ya viven en los formatos habituales de contratación tecnológica cumplen el 43,5 %; los específicos de sistemas algorítmicos, el 13,9 %; ninguno de los primeros está en cero y **los cuatro ceros absolutos están todos entre los segundos**, con una diferencia que no es atribuible al azar. Y la forma del cumplimiento por fase lo remata: el máximo está en el contrato, que es donde hay formatos, y los dos mínimos están en la preparación y en la operación, que es donde no hay ninguno. **El problema de la compra pública de IA en la región no es de convicción ni de capacidad: es que no existe el campo donde escribir el requisito.**

Esta edición también corrigió dos cosas propias. Una es aritmética y menor en consecuencias, mayor en lo que revela: el marco decía calibrarse con tres ceros y la medición tenía cuatro. El cuarto, la regla de ambición sobre madurez del proveedor, queda ahora resuelto como campo obligatorio de justificación en la declaración, con el argumento de que lo que no se verifica por existencia de documento no puede ser condición de entrada.

La otra corrección es mayor y más incómoda. El original presentaba la adopción por una agencia central de compras como la apuesta de 2026 y mencionaba la práctica chilena solo de pasada. **ChileCompra publicó Bases Tipo para licitar algoritmos e inteligencia artificial con requisitos éticos en enero de 2023**, tras consulta pública, con hitos de transparencia, privacidad, no discriminación y explicabilidad, exigencias de métricas de equidad y análisis de sesgo, y con tres instrumentos adicionales publicados como bienes públicos por el laboratorio universitario que los desarrolló con financiamiento de la banca regional. La ruta que este marco propuso como hipótesis se recorrió en la región tres años antes, y la serie la había citado como adjetivo.

La consecuencia no es que el marco sobre: es que su aporte queda medido con precisión. De los cinco estándares comparados, el MCR es el único con conformidad publicada por proceso y el único con instrumento de medición y línea de base regional. Y es también **el único que ninguna agencia central ha adoptado**. Chile tiene las bases y no tiene la conformidad publicada; Estados Unidos tiene el ciclo completo y documenta hacia adentro; Europa tiene las cláusulas y no la fase previa ni la operación. El hueco que el MCR llena es real, estrecho y verificable — y su casilla vacía también.

Lo que queda por delante no es un problema de diseño. La mediana del 27,9 % es el precio anual del statu quo, medido ítem por ítem, y su composición es conocida: cero clasificaciones de riesgo, cero fichas públicas, cero métricas publicadas, una auditoría en doce, tres renovaciones evaluadas en once. El costo de corregirlo por la ruta de entidad es una resolución y una plantilla; por la ruta

de agencia central, un ejercicio normativo ordinario con precedente regional probado. **Y la venta para corregir un contrato es su firma**, no la medición que vendrá después.

Esta edición deja tres tareas para sus propios autores: someter el marco a consulta pública, como Chile hizo con sus bases; publicar el corte de datos que sostiene una de sus tres calibraciones y que no es reproducible; e incluir procesos chilenos en la próxima medición, para saber qué produjeron unas bases tipo vigentes desde 2023 frente a una mediana colombiana del 27,9 %. Un estándar que exige a los gobiernos demostrar lo que afirman no puede pedir que se le crea.

Anexos

Anexo A. Texto modelo de la declaración de conformidad

Una página, publicada con el expediente del proceso. Los campos entre corchetes se diligencian. **No es un concepto jurídico:** cada entidad debe adaptarlo a su régimen contractual.

DECLARACIÓN DE CONFORMIDAD CON EL MARCO DE COMPRA RESPONSABLE DE IA

Bloque 1 · Identificación

Campo	Contenido
1. Entidad	[nombre]
2. Proceso	[número y objeto] · Canal: [licitación / selección abreviada / mínima cuantía / otro] · Valor: [monto]
3. Sistema objeto del proceso	[denominación y descripción en una frase]
4. Decisión pública que el sistema apoya o produce	[descripción concreta: «priorizar visitas de inspección», «verificar cumplimiento de requisitos del subsidio X»]
5. Clasificación de riesgo (CE1)	[nivel] · Criterio aplicado: [matriz de impacto de la serie / otro, identificarlo] · Documento donde consta: [enlace]
6. Ficha en el registro (CE2)	[enlace a la ficha pública del sistema]
7. Justificación de la regla de ambición	[dos o tres frases: por qué la madurez de la solución elegida es proporcional a la complejidad del problema, y qué alternativa de menor sofisticación se consideró y descartó]

Bloque 2 · Conformidad

Campo	Contenido
8. Variante aplicada	[estándar completo, 30 ítems / variante simplificada de bajo impacto, 12 ítems] · Fundamento: [clasificación del campo 5]
9. Ítems aplicables	[n] de [30 / 12] · Ítems no aplicables y por qué: [lista breve]
10. Ítems cumplidos	[n] de [n aplicables] = [%]
11. Métricas de operación pactadas (CE3)	Métricas: [lista] · Umbrales: [valores] · Periodicidad: [trimestral u otra] · Cláusula: [número] · Canal de publicación: [enlace]
12. Evidencia	[tabla o lista con un enlace por ítem cumplido]
13. Excepciones	[ítem, motivo, firma del ordenador del gasto o del responsable de gobernanza de IA]
14. Resultado	[CONFORME / NO CONFORME] · Un proceso es no conforme si falta cualquiera de CE1, CE2 o CE3, con independencia del porcentaje del campo 10
15. Responsables	Gestor técnico del contrato: [nombre y cargo] · Responsable de gobernanza de IA: [nombre y cargo]
16. Fecha y versión del marco	[fecha] · MCR v2

Notas de diseño de la declaración.

- **El campo 14 es el que da sentido al instrumento.** Una declaración que puede decir «no conforme» es información; una que solo puede decir «conforme» es propaganda. Las entidades deben publicar declaraciones no conformes: es lo que hace creíbles las conformes.
- **El campo 7 es la respuesta al cuarto cero (§8.2).** Es el único campo que exige argumentar en lugar de enlazar, y por eso es el más fácil de llenar mal y el más útil de leer.
- **El campo 12 es lo que la verificación audita (Anexo D).** Enlaces que no abren o que no documentan el ítem son incumplimiento con agravante.
- **El campo 9 es la defensa contra la acusación de rigidez.** Declarar qué no aplica y por qué es lo que permite que el estándar sea exigente sin ser absurdo.

Anexo B. Clasificación de los treinta ítems (G1 / G2) y datos del análisis

Sostiene el análisis de la §11.3 y se publica para que pueda disputarse y el cálculo repetirse con otra codificación.

Criterio de clasificación, aplicado a cada ítem: ¿el requisito existe ya, en alguna forma, en los formatos habituales de contratación tecnológica del Estado? **G1** = sí (requisito genérico). **G2** = no (requisito específico de sistemas algorítmicos).

Fase	Ítem	Grupo	Cumplimiento	%
Fase 2	Niveles de servicio sobre métricas	G1	14 de 20	70,0 %
Fase 1	Seguridad y continuidad de terceros	G1	13 de 20	65,0 %
Fase 1	Documentación técnica exigida	G1	12 de 20	60,0 %
Fase 2	Hitos de validación verificables	G1	10 de 20	50,0 %
Fase 0	Base legal de datos verificada	G1	9 de 20	45,0 %
Fase 1	Criterios con perfil técnico en comité	G1	9 de 20	45,0 %
Fase 1	Objeto por resultados y métricas	G1	8 de 20	40,0 %
Fase 3	Gestor de contrato técnico	G1	8 de 20	40,0 %
Fase 0	Diálogo de mercado sobre el problema	G1	5 de 14	35,7 %
Fase 0	Problema y resultados por escrito	G1	7 de 20	35,0 %
Fase 2	Propiedad y exportación de datos clara	G1	7 de 20	35,0 %
Fase 3	Aceptación contra pruebas pactadas	G1	6 de 20	30,0 %
Fase 1	Costo total de vida útil exigido	G1	3 de 20	15,0 %
Fase 4	Cierre con exportación verificada	G2	2 de 6	33,3 %
Fase 2	Diez cláusulas mínimas (núcleo)	G2	6 de 20	30,0 %
Fase 1	Prueba práctica como criterio	G2	4 de 14	28,6 %
Fase 4	Renovación con evaluación documentada	G2	3 de 11	27,3 %
Fase 3	Protocolo de supervisión aprobado	G2	3 de 14	21,4 %
Fase 2	Plan de salida con transición	G2	4 de 20	20,0 %
Fase 0	Alternativas no algorítmicas evaluadas	G2	3 de 20	15,0 %
Fase 1	Ficha de datos exigida	G2	2 de 14	14,3 %
Fase 1	Desempeño por grupos exigido	G2	1 de 9	11,1 %
Fase 2	Gestión de cambios de modelo	G2	2 de 20	10,0 %
Fase 3	Línea de base medida antes de operar	G2	2 de 20	10,0 %
Fase 4	Auditoría externa realizada	G2	1 de 12	8,3 %
Fase 4	Monitoreo de deriva con umbrales	G2	1 de 14	7,1 %
Fase 0	Clasificación en matriz de impacto	G2	0 de 20	0,0 %
Fase 0	Madurez contrastada (regla de ambición)	G2	0 de 20	0,0 %
Fase 3	Ficha publicada en registro	G2	0 de 20	0,0 %
Fase 4	Métricas trimestrales publicadas	G2	0 de 17	0,0 %

Tabla 15. Clasificación completa de los treinta ítems. Fuente: clasificación de Enpath sobre el scorecard de noviembre de 2025.

Estadísticos del análisis. G1: n = 13, media 43,5 %, mediana 40,0 %, ceros 0. G2: n = 17, media 13,9 %, mediana 11,1 %, ceros 4. Diferencia de medias: 29,6 puntos porcentuales. Prueba de permutación de la diferencia de medias, 200.000 remuestreos, una cola: $p \approx 0,00001$.

Los tres ítems cuya clasificación es más discutible, declarados para el escrutinio: - «**Base legal de datos verificada**» (G1, 45 %): se clasificó como genérico porque la verificación de base legal para tratar datos personales es exigencia de cualquier contrato que involucre datos. Un revisor podría argumentar que en IA tiene especificidad propia. - «**Diez cláusulas mínimas (núcleo)**» (G2, 30 %): se clasificó como específico porque las diez cláusulas son la aportación de esta serie para contratos de IA. Un revisor podría notar que varias de ellas son adaptaciones de cláusulas convencionales. - «**Costo total de vida útil exigido**» (G1, 15 %): es el ítem de G1 con peor cumplimiento y el que más tensiona la clasificación. Se mantuvo en G1 porque el costo total de propiedad es una exigencia estándar de compra de tecnología.

Prueba de robustez: mover los tres ítems al grupo opuesto reduce la diferencia pero no la elimina; el resultado no depende de ninguno de ellos por separado. Cualquier lector puede recalcular con el archivo publicado del scorecard y el script `data/charts_report.py`.

Anexo C. Criterios de verificación de las tres condiciones de entrada

Cada condición se verifica de forma **binaria**, por existencia de documento localizable. La duda se resuelve contra el proceso.

CE1 · Clasificación de riesgo

- **Se acepta:** documento del expediente que asigne al sistema un nivel de riesgo o impacto, identificando el criterio o matriz aplicada y el fundamento de la asignación.
- **No se acepta:** mención genérica del riesgo del proyecto; matriz de riesgos contractuales; afirmación de que el sistema «no representa riesgos» sin criterio declarado.
- **Fundamento externo:** el test de proporcionalidad en sentido estricto de la doctrina colombiana de protección de datos aplicable a IA⁶; la graduación por riesgo del régimen europeo².

CE2 · Ficha en el registro

- **Se acepta:** ficha pública del sistema con propósito, decisión que apoya, datos que usa, forma de supervisión humana y responsable, enlazada desde el contrato o desde el expediente.
- **No se acepta:** descripción comercial del proveedor; nota de prensa; mención del sistema en un informe de gestión sin ficha.
- **Fundamento externo:** los registros neerlandés y británico y el registro federal canadiense de sistemas de decisión automatizada.

CE3 · Métricas de operación

- **Se acepta:** cláusula contractual que identifique las métricas de desempeño y supervisión, sus umbrales y su periodicidad, y que comprometa su publicación.
- **No se acepta:** niveles de servicio de disponibilidad técnica sin métricas de calidad de la decisión; compromiso de «reportar» sin definir qué ni con qué periodicidad; métricas sin umbral.
- **Nota:** es la condición más exigente y la que ninguna jurisdicción examinada exige. Se conserva porque es la única que gobierna la fase donde el cumplimiento medido es del 15,2 %.

Anexo D. Protocolo de verificación externa por muestreo

Ejecutable por un órgano de control, un par institucional o un tercero independiente. No requiere acceso privilegiado.

1. Universo y muestra. Todas las declaraciones de conformidad publicadas por la entidad en el periodo. Muestra mínima: la mayor entre cinco declaraciones o el 20 % del total, priorizando procesos con clasificación de riesgo alta.

2. Regla central. La verificación se hace contra la evidencia enlazada, no contra la declaración. Una declaración que afirma cumplimiento con enlaces que no abren, no corresponden al ítem o no documentan lo afirmado constituye incumplimiento con agravante, y así debe reportarse.

3. Secuencia por declaración. 1. Verificar las tres condiciones de entrada con los criterios del Anexo C. Si falta una, el proceso es no conforme y el campo 14 de la declaración debe decirlo; si dice «conforme», se registra como **declaración falsa**. 2. Verificar una muestra aleatoria de cinco ítems del campo 12 contra su evidencia enlazada. 3. Verificar que el campo 9 (ítems no aplicables) sea consistente con la variante declarada en el campo 8 y con la clasificación del campo 5. 4. Leer el campo 7 (regla de ambición) y valorar si contiene un argumento o una fórmula vacía. **Es el único juicio no binario del protocolo, y debe reportarse como observación, no como incumplimiento.**

4. Producto. Tabla con tres columnas por declaración: condiciones de entrada verificadas; ítems verificados y confirmados; discrepancias encontradas. Y una conclusión de una línea por declaración: **confirmada, confirmada con observaciones, o desmentida.**

5. Publicación. El informe de verificación se publica. Una verificación no publicada reproduce el problema que el marco intenta resolver.

6. Periodicidad. Anual, con rotación de entidades.

Anexo E. Glosario

Término	Definición operativa en este documento
MCR	Marco Enpath de compra responsable de IA: cinco principios, estándar v2 de treinta ítems y tres piezas de conformidad
Condición de entrada (CE)	Requisito sin el cual el proceso no se considera conforme, con independencia del cumplimiento de los demás ítems
Regla de canal	El estándar aplica por objeto — sistema que decide o apoya decisiones — y no por modalidad, cuantía o canal de contratación
Graduación	Aplicación del estándar en variante simplificada de doce ítems para sistemas de bajo impacto
Regla de ambición	Exigencia de contrastar la madurez de la solución y del proveedor con la complejidad del problema
Declaración de conformidad	Documento de una página, publicado con el expediente, que declara qué ítems aplicaban, cuáles se cumplen y dónde está la evidencia
Verificación por muestreo	Auditoría externa de una muestra de declaraciones publicadas contra su evidencia enlazada (Anexo D)
G1 / G2	Clasificación de los treinta ítems según si el requisito existe o no en los formatos habituales de contratación tecnológica (Anexo B)
Asimetría documental	Situación en que la industria global produce, por obligación de otra jurisdicción, documentación que la demanda regional no exige
Conformidad de papel	Declaración publicada cuya evidencia enlazada no sostiene lo declarado

Anexo F. Índice de figuras y tablas

Figuras 1. Los tres niveles del MCR (§7) 2. Las tres rutas de adopción (§16) 2. Los treinta ítems ordenados por cumplimiento (§11.1) 3. El descenso por fase (§11.2) 4. La hipótesis de los formularios, puesta a prueba (§11.3) 5. Los cuatro ceros y las tres condiciones de entrada (§8.1) 6. El MCR frente a sus pares (§12) 7. Las tres rutas, en detalle (§16) 8. Matriz de riesgos del marco (§14)

Tablas 1. La línea de base del estándar, recomputada (§1.1) 2. La trayectoria de la agenda de compra (§4) 3. Las ocho causas y su huella medida (§5.1) 4. El marco aplicable frente a las tres condiciones de entrada (§6) 5. Las tres condiciones de entrada (§8.1) 6. Cumplimiento por fase (§11.2) 7. La prueba de la hipótesis de los formularios (§11.3) 8. Denominadores de los treinta ítems (§11.5) 9. El MCR frente a cuatro pares (§12) 10. El caso chileno frente a los elementos del MCR (§12.4) 11. Costo, plazo y alcance de las tres rutas (§13.1) 12. Análisis de actores (§13.2) 13. Registro de riesgos (§14) 14. Predicciones falsables (§17) 15. Escala de calidad de la evidencia (§18.3) 16. Clasificación completa de los treinta ítems (Anexo B)

Referencias

Fuentes primarias colombianas

1. Superintendencia de Industria y Comercio (2024). *Circular Externa 002 del 21 de agosto de 2024: lineamientos sobre el tratamiento de datos personales en sistemas de inteligencia artificial*. <https://sedeelectronica.sic.gov.co/sites/default/files/normativa/Circular%20Externa%20No.%20002%20del%2021%20de%20agosto%20de%202024.pdf>
2. Departamento Nacional de Planeación (2025). *Documento CONPES 4144: Política Nacional de Inteligencia Artificial*, aprobado el 14 de febrero de 2025. <https://www.dnp.gov.co/publicaciones/Planeacion/Paginas/conpes-4144-hoja-de-ruta-colombia-inteligencia-artificial-retos-actuales-transformacion-futura.aspx>
3. Congreso de la República de Colombia (2014). *Ley 1712 de 2014: transparencia y derecho de acceso a la información pública*. <https://www.funcionpublica.gov.co/eva/gestornormativo/norma.php?i=56882>
4. Congreso de la República de Colombia (2012). *Ley 1581 de 2012*; y (2008) *Ley 1266 de 2008*.
5. Agencia Nacional de Contratación Pública — Colombia Compra Eficiente. *Pliegos y documentos tipo y Boletín «Radar de la compra pública», edición 3 de 2026* (que incluye orientación sobre contratación pública de soluciones tecnológicas con inteligencia artificial). <https://www.colombiacompra.gov.co/archivos/tipo/pliegos-tipo>

Fuentes primarias internacionales

1. Comisión Europea, Community of Practice on Public Procurement of AI (2025). *Updated EU AI Model Contractual Clauses (MCC-AI)*, 5 de marzo de 2025. <https://public-buyers-community.ec.europa.eu/communities/procurement-ai/resources/updated-eu-ai-model-contractual-clauses>
2. Unión Europea (2024). *Reglamento (UE) 2024/1689 (Reglamento de Inteligencia Artificial)*. <https://eur-lex.europa.eu/eli/reg/2024/1689/oj>
3. Unión Europea (2026). *Digital Omnibus on AI*: adopción del Parlamento Europeo el 16 de junio de 2026; adopción del Consejo el 29 de junio de 2026; firma del acto el 8 de julio de 2026.
4. Office of Management and Budget, Estados Unidos (2025). *Memorandum M-25-22: Driving Efficient Acquisition of Artificial Intelligence in Government*, 3 de abril de 2025. <https://www.whitehouse.gov/wp-content/uploads/2025/02/M-25-22-Driving-Efficient-Acquisition-of-Artificial-Intelligence-in-Government.pdf>
5. ChileCompra (2023). *Bases Tipo para la licitación de proyectos de algoritmos e inteligencia artificial con requisitos éticos*. <https://www.chilecompra.cl/2023/01/ya-se-encuentra-disponible-bases-tipo-para-licitar-proyectos-de-algoritmos-e-inteligencia-artificial-con-requisitos-eticos/>
6. GobLab, Universidad Adolfo Ibáñez. *Proyecto Algoritmos Éticos, Responsables y Transparentes: Evaluación de Impacto Algorítmico, medición de sesgo y equidad, y Ficha de Transparencia Algorítmica*. <https://goblab.uai.cl/>

7. World Economic Forum (2020). *AI Procurement in a Box*: guías de adquisición, cuaderno de trabajo y lecciones de implementación, 11 de junio de 2020. <https://www.weforum.org/publications/ai-procurement-in-a-box/>
8. Gobierno de los Países Bajos. *Het Algoritmeregister*. <https://algoritmes.overheid.nl/en> · Gobierno del Reino Unido. *Algorithmic Transparency Recording Standard Hub*. <https://www.gov.uk/government/collections/algorithmic-transparency-recording-standard-hub> · Treasury Board of Canada Secretariat. *Directive on Automated Decision-Making*. <https://www.tbs-sct.canada.ca/pol/doc-eng.aspx?id=32592>

Serie de investigación de Enpath

1. Enpath (2024). *Compra pública de IA: las ocho causas y las diez cláusulas mínimas* (agosto de 2024). Esta biblioteca.
2. Enpath (2025). *Estándares de compra pública para sistemas de IA: treinta ítems en cinco fases* (mayo de 2025); *Scorecard de compra pública de IA: veinte procesos* (noviembre de 2025, con archivo de datos resultados_scorecard.csv); *Marco de gobernanza ministerial* (enero de 2025); *Plantilla de ficha de registro de algoritmos* (abril de 2025); *Evaluación específica de riesgo* (septiembre de 2025). Esta biblioteca.
3. Enpath (2026). *Benchmark de capacidades de IA en el gobierno* (febrero de 2026); *Índice Enpath de gobernanza de IA del sector público: Colombia 2026* (marzo de 2026); *Estándares de transparencia en la compra de IA: las cinco ventanas* (mayo de 2026). Esta biblioteca.

Notas

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

1. ChileCompra (Dirección de Compras y Contratación Pública de Chile) y GobLab de la Universidad Adolfo Ibáñez. *Bases Tipo para la licitación de proyectos de algoritmos e inteligencia artificial con requisitos éticos*, disponibles desde enero de 2023, tras consulta pública del 5 de octubre al 18 de noviembre de 2022 con respuestas publicadas el 2 de diciembre de 2022. Incorporan hitos de transparencia, privacidad, no discriminación y explicabilidad, y exigen —según la naturaleza del contrato— modelos con métricas estadísticas de equidad, medidas adicionales de protección de datos y análisis de sesgo. Desarrolladas en el proyecto «Algoritmos Éticos, Responsables y Transparentes», ejecutado por el GobLab UAI con apoyo de BID Lab desde diciembre de 2020, con ChileCompra, el Ministerio de Ciencia, Tecnología, Conocimiento e Innovación, la División de Gobierno Digital de Segpres, el Consejo para la Transparencia y varias entidades prestadoras como socios. El proyecto publicó además, entre 2023 y 2025, la Evaluación de Impacto Algorítmico, la herramienta de medición de sesgo y equidad y la Ficha de Transparencia Algorítmica como bienes públicos; acumulaba cuatro reconocimientos internacionales a julio de 2024; y ChileCompra aplicó la evaluación de impacto a su propio proyecto de detección de anomalías en procesos de compra en diciembre de 2025.
2. Comisión Europea, *Model Contractual Clauses for AI Procurement* (MCC-AI), versión actualizada publicada el 5 de marzo de 2025 por la Community of Practice on Public Procurement of AI: versión completa para IA de alto riesgo, versión ligera

para IA no de alto riesgo y comentario de uso. Siguen en lo sustancial las obligaciones del capítulo III del Reglamento (UE) 2024/1689. Sustituye a la versión piloto de 2023.

3. Office of Management and Budget de Estados Unidos, *Memorandum M-25-22: Driving Efficient Acquisition of Artificial Intelligence in Government*, del 3 de abril de 2025, que deroga y reemplaza el M-24-18. Tres ejes: garantizar un mercado de IA competitivo evitando la dependencia de proveedor y promoviendo la interoperabilidad; proteger el gasto público mediante seguimiento del desempeño de la IA y gestión de riesgos; y promover la adquisición efectiva mediante equipos interfuncionales. Cubre el ciclo completo de adquisición, desde la identificación de requisitos y el estudio de mercado hasta la selección, la administración del contrato y su cierre; exige validación de desempeño y pruebas previas al fallo para IA de alto impacto; protege derechos sobre datos y propiedad intelectual; plazo de cumplimiento pleno de 180 días.
4. *Digital Omnibus on AI* de la Unión Europea. Acuerdo provisional el 7 de mayo de 2026; adopción por el Parlamento Europeo el 16 de junio de 2026; adopción por el Consejo el 29 de junio de 2026; firma del acto el 8 de julio de 2026. Traslada las obligaciones de los sistemas de alto riesgo del Anexo III del 2 de agosto de 2026 al 2 de diciembre de 2027, y las del Anexo I al 2 de agosto de 2028; mantiene las obligaciones de transparencia del artículo 50 desde el 2 de agosto de 2026, con gracia hasta el 2 de diciembre de 2026 para el marcado de contenido de sistemas ya en el mercado; conserva el deber de registro para sistemas autoclasificados como no de alto riesgo; y posterga al 2 de agosto de 2027 el plazo para establecer espacios controlados de pruebas. Las obligaciones de los modelos de propósito general son aplicables desde agosto de 2025.
5. World Economic Forum, *AI Procurement in a Box*, publicado el 11 de junio de 2020: incluye guías de adquisición de IA para gobiernos, un cuaderno de trabajo para funcionarios de política y de compras, y un documento de lecciones y oportunidades durante la implementación. Desarrollado por la red global del Centro para la Cuarta Revolución Industrial, con representantes de los sectores público y privado, la academia y la sociedad civil. Pilotado con departamentos y agencias de Bahrein y el Reino Unido y con la empresa de electricidad y agua de Dubái, con retroalimentación de pilotos en Colombia y otros gobiernos socios; concebido como documento vivo.
6. Superintendencia de Industria y Comercio de Colombia, Circular Externa 002 del 21 de agosto de 2024: lineamientos sobre el tratamiento de datos personales en sistemas de inteligencia artificial. Exige que el tratamiento supere los criterios de idoneidad, necesidad, razonabilidad y proporcionalidad en sentido estricto; recomienda estudios de impacto de privacidad; y desarrolla los principios de transparencia, privacidad, seguridad, responsabilidad demostrada y protección de los derechos de niños, niñas y adolescentes, en el marco de las Leyes 1266 de 2008 y 1581 de 2012.
7. Departamento Nacional de Planeación de Colombia, *Documento CONPES 4144: Política Nacional de Inteligencia Artificial*, aprobado el 14 de febrero de 2025: seis ejes estratégicos, más de cien acciones con horizonte 2030 e inversión indicativa de 479.000 millones de pesos; entidades líderes MinTIC, DAPRE, MinCiencias y DNP.