Enpath

NOTA TÉCNICA · INVESTIGACIÓN

Gobernanza de datos para instituciones públicas: un marco práctico

Mayo de 2024

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

Enpath · Investigación · Marco (framework) · Pilar: Gobernanza de datos **Publicado:** mayo de 2024 · **Corte de información:** 31 de mayo de 2024 · **Idioma:** español

> Ninguna entidad pública tiene un problema de inteligencia artificial que no sea, primero, un problema de datos. Este documento propone un marco de gobernanza de datos implementable por cualquier institución pública de la región con las capacidades que ya tiene: cuatro capas, diez componentes y una ruta de 90 días.

Actualización.

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

- Chile promulgó en 2024 su nueva ley de protección de datos personales, con una agencia de protección de datos; era el principal rezago normativo señalado en la sección regional de este informe.
- El Reglamento de IA europeo entró en vigor en agosto de 2024; sus requisitos de gobernanza de datos para sistemas de alto riesgo (art. 10) confirmaron la tesis central de este documento: la calidad y trazabilidad de los datos se volvió requisito legal, no buena práctica.
- En Colombia, la Superintendencia de Industria y Comercio publicó en 2024 lineamientos sobre tratamiento de datos personales en sistemas de inteligencia artificial, y el CONPES 4144 (febrero de 2025) incluyó la infraestructura de datos como habilitador de la política nacional de IA.

Resumen ejecutivo.

La gobernanza de datos suele presentarse como un problema técnico de sistemas de información. Es, ante todo, un problema de gestión pública: **saber qué datos tiene la entidad, quién responde por ellos, con qué calidad, quién puede usarlos y bajo qué reglas.** Una entidad que no puede responder esas cinco preguntas no puede cumplir bien su misión ordinaria — y no debería, con esos cimientos, adoptar inteligencia artificial.

El momento de este documento no es casual. El Consejo de la Unión Europea adoptó formalmente el Reglamento de IA el 21 de mayo de 2024; entre sus requisitos para sistemas de alto riesgo está precisamente la gobernanza de los datos de entrenamiento y operación. La banca multilateral condiciona cada vez más su financiamiento digital a capacidades de datos. Y la región acumula recordatorios de lo que cuesta no gobernarlos: el ataque de ransomware a un proveedor tecnológico en septiembre de 2023 interrumpió servicios de decenas de entidades públicas colombianas, incluida la rama judicial, exponiendo la dependencia sin salvaguardas de infraestructuras de terceros.

Colombia tiene una posición de partida mejor que la mayoría: el marco normativo está razonablemente completo —habeas data (2008), protección de datos (2012), transparencia (2014), política de explotación de datos (2018, la primera de la región), infraestructura de datos (2021)— y el país se ha destacado en datos abiertos, con el tercer lugar mundial en el índice OURdata de la OCDE en 2019 y un desempeño por encima del promedio OCDE en la edición 2023 (OCDE, 2023). La brecha no es de normas ni de vitrinas: es de gestión interna. La mayoría de las entidades no tiene inventario de sus datos, ni responsable identificable, ni reglas de calidad y compartición.

El **Marco Enpath de gobernanza de datos** —cuatro capas, diez componentes— traduce los estándares internacionales a la escala de una entidad pública latinoamericana, y la **ruta de 90 días** define por dónde empezar sin esperar presupuesto extraordinario.

Hallazgos principales.

1. **El marco normativo colombiano está sustancialmente completo; la implementación institucional no.** Entre 2008 y 2022 el país expidió las piezas normativas esenciales. Ninguna norma, sin embargo, obliga a cada entidad a mantener un inventario de datos con responsable designado — el componente sin el cual las demás obligaciones son inaplicables.
2. **Los datos son ya la condición de entrada a la agenda de IA.** El Reglamento de IA europeo (adoptado el 21 de mayo de 2024) exige a los sistemas de alto riesgo prácticas de gobernanza sobre los datos: pertinencia, representatividad, gestión de sesgos, trazabilidad. Quien no gobierna sus datos no podrá comprar ni operar sistemas conformes.
3. **La región tiene liderazgo en vitrina y rezago en trastienda.** Colombia, Brasil y Uruguay destacan en datos abiertos y plataformas de interoperabilidad; la gestión interna —calidad, catálogos, ciclo de vida— permanece débil y no se mide. Los índices internacionales evalúan lo primero, casi nunca lo segundo.
4. **La seguridad es el componente donde el costo de la omisión ya se materializó.** El incidente de septiembre de 2023 en Colombia mostró que la continuidad de servicios públicos depende de proveedores cuya gestión de riesgos las entidades no verifican. La gobernanza de datos incluye gobernar los contratos que los custodian.
5. **Uruguay demuestra que la escala no es excusa.** Con una agencia rectora (AGESIC), una plataforma de interoperabilidad operativa y protección de datos con reconocimiento de adecuación europea desde 2012, Uruguay opera el sistema de datos públicos más coherente de la región — con una fracción del presupuesto de sus vecinos grandes.

Por qué importa.

Tres decisiones concretas dependen de este marco. Para la dirección de una entidad: **qué debe existir en su organización antes de firmar cualquier proyecto de analítica o IA** — y qué exigir a quien se lo proponga. Para las oficinas de planeación y TI: **cómo pasar de proyectos de datos aislados a una operación gobernada**, con roles y reglas que sobrevivan a los contratistas. Para los órganos de control: **qué preguntas hacer** para distinguir una entidad que gestiona datos de una que solo los

acumula. La gobernanza de datos no es un proyecto de tecnología; es la condición para que los proyectos de tecnología no fracasen.

Antecedentes.

La discusión internacional maduró en una década. La OCDE formuló en 2019 el modelo del "sector público conducido por datos": los datos como activo estratégico que exige liderazgo, arquitectura y controles, no solo portales (OCDE, 2019). La profesión consolidó sus estándares (el cuerpo de conocimiento DAMA-DMBOK es la referencia de facto para la gestión de datos organizacional). Y la Unión Europea convirtió la agenda en derecho positivo: el Reglamento de Gobernanza de Datos —aplicable desde septiembre de 2023— y el Reglamento de Datos —en vigor desde enero de 2024— crean las instituciones del intercambio confiable de datos; el Reglamento de IA cierra el ciclo exigiendo gobernanza de datos como requisito de los sistemas de alto riesgo (Unión Europea, 2022; 2024).

América Latina siguió una trayectoria propia: pionera en datos abiertos y leyes de protección de datos (Argentina 2000, Uruguay 2008 — ambas con reconocimiento de adecuación de la UE), y más lenta en la infraestructura y la gestión. Colombia ilustra ambos lados.

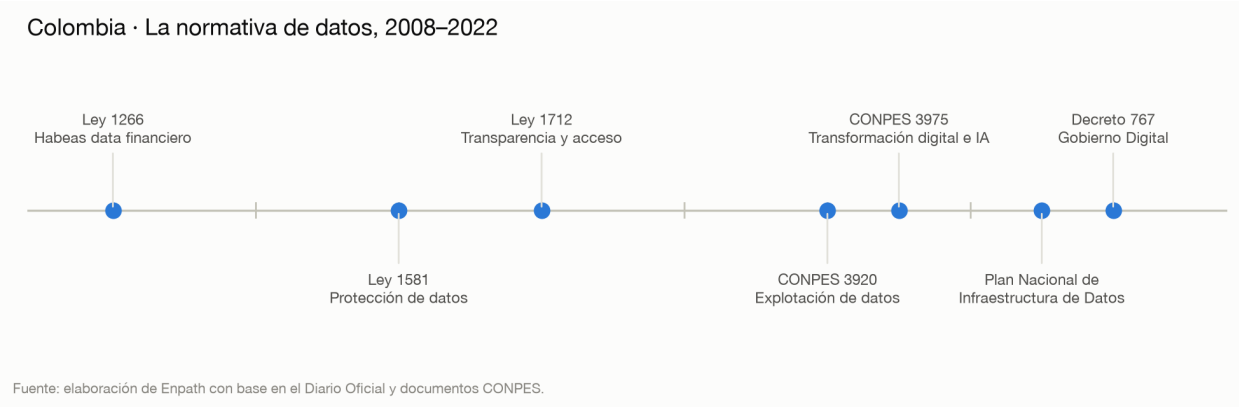

Figura 1. Colombia: hitos normativos de datos, 2008–2022. Fuente: elaboración de Enpath.

El marco.

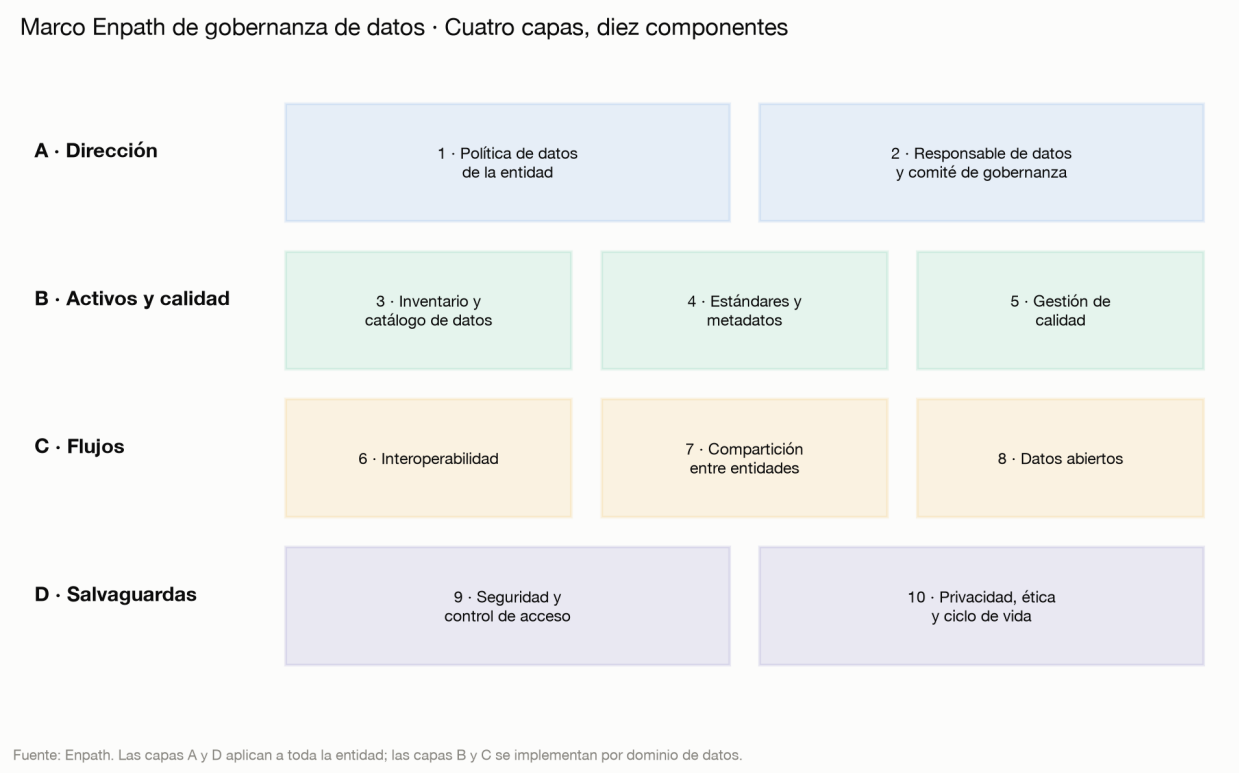

Figura 2. Marco Enpath de gobernanza de datos. Fuente: Enpath.

Capa A · Dirección.

1 · Política de datos de la entidad. Un documento corto, aprobado por la dirección, que declara: qué dominios de datos gestiona la entidad, con qué fines, bajo qué normas y con qué prioridades. No es un manual técnico; es el mandato.

2 · Responsable de datos y comité. Una persona del nivel directivo responde por los datos de la entidad (en la práctica internacional, el *chief data officer*); un comité pequeño —jurídica, TI, planeación, misional— resuelve los conflictos que los datos siempre generan: quién accede, qué se comparte, qué se publica. *Señal de alerta:* si la respuesta a "¿quién responde por los datos?" es "sistemas", la entidad no tiene gobernanza; tiene infraestructura.

Capa B · Activos y calidad.

3 · Inventario y catálogo. La pieza fundacional: qué bases de datos existen, qué contienen, quién es su dueño funcional, qué base legal las ampara y qué calidad tienen. Sin inventario no hay seguridad (no se protege lo que no se conoce), ni interoperabilidad, ni cumplimiento.

4 · Estándares y metadatos. Identificadores comunes (persona, empresa, predio, territorio), formatos y diccionarios de datos. La lección regional: cada identificador no estandarizado se paga después en cada cruce de bases.

5 · Gestión de calidad. Reglas de validación en origen, medición periódica (completitud, unicidad, vigencia) y responsables de corrección por dominio. La calidad no se audita una vez: se opera.

Capa C · Flujos.

6 · Interoperabilidad. Conexión técnica y semántica con la plataforma estatal (en Colombia, los servicios ciudadanos digitales y el modelo adoptado de la experiencia estonia de X-Road). El principio rector: el ciudadano no es el mensajero del Estado; los datos que el Estado ya tiene no se le vuelven a pedir.

7 · Compartición entre entidades. Acuerdos escritos de intercambio: finalidad, base legal, campos, vigencia, registro de accesos. Este componente se desarrolla en profundidad en un informe posterior de esta serie.

8 · Datos abiertos. Publicación por defecto de lo no reservado, con calidad de reutilización. Colombia tiene aquí su fortaleza histórica —tercer lugar mundial en el índice OURdata 2019; sobre el promedio OCDE en 2023— que conviene mantener, pero que no sustituye las capas internas (OCDE, 2023).

Capa D · Salvaguardas.

9 · Seguridad y control de acceso. Clasificación de la información, mínimo privilegio, registro de accesos, y —la lección de 2023— gestión de riesgos de terceros: continuidad, respaldo y auditoría sobre los proveedores que custodian datos o infraestructura.

10 · Privacidad, ética y ciclo de vida. Cumplimiento del régimen de protección de datos (Ley 1581 de 2012 y jurisprudencia), evaluación de impacto para tratamientos riesgosos, anonimización para publicación, y disciplina de archivo y eliminación (tablas de retención documental).

Figura 3. Ciclo de vida del dato público y controles mínimos. Fuente: Enpath.

Situación actual.

El diagnóstico en tres planos.

Normativo: completo en lo esencial. Colombia no necesita una gran ley nueva para gobernar sus datos; necesita cumplir las que tiene con instrumentos de gestión.

Institucional: débil. No existe la figura obligatoria de responsable de datos por entidad; los inventarios son excepcionales; la calidad no se mide ni se reporta. El Plan Nacional de Infraestructura de Datos

(2021) formuló la arquitectura correcta; su adopción entidad por entidad avanza sin mecanismo de exigencia (MinTIC, 2021).

Operativo: heterogéneo. Conviven islas de excelencia —el DANE y su sistema estadístico, la infraestructura de datos abiertos (datos.gov.co), operaciones de interoperabilidad en crecimiento— con la norma general: bases duplicadas, identificadores inconsistentes, dependencia de contratistas y conocimiento no documentado.

Análisis de datos.

La evidencia comparada disponible converge en el mismo patrón:

- **En datos abiertos, la región compite arriba.** El índice OURdata 2023 de la OCDE sitúa a Corea y Francia a la cabeza; Colombia aparece en el grupo que puntúa por encima del promedio OCDE — un resultado notable para el único país andino de la organización (OCDE, 2023).
- **En gobernanza interna, la evidencia es indirecta pero consistente.** El estudio OCDE–CAF sobre IA en el sector público regional identificó la gobernanza de datos como la debilidad transversal de los proyectos: datos fragmentados, sin calidad medida y sin reglas de compartición (OCDE y CAF, 2022). El ILIA 2023 confirma el contraste entre disponibilidad de datos y capacidad de uso (CENIA y CEPAL, 2023).
- **En seguridad, el costo ya es observable.** El incidente de ransomware de septiembre de 2023 en Colombia —que afectó la operación de decenas de entidades a través de un solo proveedor— es el caso de estudio regional de riesgo de terceros no gobernado.

Pregunta de control	Si la respuesta es "no"
¿Existe un inventario de bases de datos con dueño por cada una?	Nada de lo demás es verificable
¿Hay un responsable de datos del nivel directivo?	La gobernanza vive en los contratistas
¿Se mide la calidad de los dominios críticos?	Las decisiones heredan errores invisibles
¿Hay acuerdos escritos para cada intercambio de datos?	Cada cruce es un riesgo jurídico
¿Se auditan los proveedores que custodian datos?	La continuidad depende de la suerte
¿Existen tablas de retención aplicadas?	El pasivo de datos crece sin límite

Tabla 1. Seis preguntas de control para una entidad pública. Fuente: Enpath.

Colombia.

El activo diferencial colombiano es doble: un sistema estadístico nacional sólido (DANE) y una tradición de datos abiertos reconocida internacionalmente. El pasivo es la gestión interna en el resto de la administración. Tres frentes concentran el trabajo pendiente:

- **Identificadores e interoperabilidad.** El país avanzó en la plataforma (servicios ciudadanos digitales, interoperabilidad con referencia al modelo estonio); el cuello de botella es la adopción por entidad y la disciplina de identificadores comunes en las bases misionales.
- **Registros administrativos para decisión.** La agenda del informe de abril de esta serie: los registros de salud, educación, protección social y tributarios permanecen fragmentados; su integración gobernada —con las salvaguardas de la capa D— es la inversión de mayor retorno del Estado colombiano en materia de datos.
- **Seguridad y terceros.** Tras el incidente de 2023, la pregunta de control mínima para toda entidad es verificable: ¿qué pasa con la operación si el proveedor principal se detiene mañana? Respaldo, continuidad y cláusulas de auditoría dejan de ser opcionales.

La ruta de 90 días (recomendaciones) está escrita para la entidad colombiana promedio: sin presupuesto nuevo, con las obligaciones normativas ya vigentes como palanca.

América Latina.

- **Uruguay** es el referente de coherencia: agencia rectora (AGESIC), plataforma de interoperabilidad en operación, protección de datos con adecuación europea desde 2012. La lección: una arquitectura pequeña y gobernada supera a una grande y dispersa (AGESIC).
- **Brasil** combina la ley de protección de datos más desarrollada de la región (LGPD, con autoridad propia, la ANPD) con la experiencia aleccionadora del decreto de compartición amplia de bases federales (2019), criticado por crear un registro ciudadano unificado sin salvaguardas proporcionales: compartir más no es gobernar mejor.
- **México, Chile, Perú y Argentina** muestran las variantes del rezago: leyes de datos personales en actualización (Chile tramitaba en 2024 su nueva ley; Argentina mantiene su adecuación europea sobre una ley de 2000), portales de datos abiertos maduros y gestión interna sin medición.
- **Regional:** las redes existentes (Red GEALC, CEPAL, CAF, BID) han producido estándares y financiamiento; el eslabón ausente es el mismo que dentro de los países: mecanismos que hagan exigible la gestión, no solo la publicación.

Comparación internacional.

- **Estonia — X-Road.** Dos décadas de intercambio de datos con registro distribuido y el principio "solo una vez": el Estado no vuelve a pedir lo que ya sabe. La lección exportable no es el

software —que Colombia ya referencia— sino la regla de gobernanza: cada consulta queda registrada y es auditable por el ciudadano (e-Estonia).

- **Unión Europea — del derecho a la infraestructura.** La secuencia europea (protección de datos → gobernanza de datos → espacios de datos → requisitos de datos en IA) muestra hacia dónde converge la regulación que la región recibirá vía proveedores y acuerdos comerciales.
- **OCDE — el sector público conducido por datos.** El marco de 2019 ordena la discusión en tres frentes —gobernanza, arquitectura, valor— y deja el mensaje que este documento adapta: los datos públicos no son un subproducto de los trámites; son un activo que exige gestión deliberada (OCDE, 2019).

Oportunidades.

- **La palanca regulatoria ya existe.** Las entidades colombianas pueden anclar el marco en obligaciones vigentes (Ley 1581, Ley 1712, política de gobierno digital): la gobernanza de datos es la forma de cumplirlas, no una carga adicional.
- **Retorno inmediato en trámites y focalización.** Cada intercambio de datos gobernado elimina requisitos al ciudadano y errores de inclusión/exclusión en programas sociales; son resultados visibles en meses, no años.
- **Preparación para la IA a costo hundido.** Las diez componentes del marco son exactamente lo que exigirá cualquier adopción seria de IA (y lo que el reglamento europeo ya exige a los sistemas de alto riesgo): invertir hoy en datos es comprar barata la opción de la IA mañana.
- **Liderazgo regional disponible.** Ningún país de la región mide la gestión interna de datos de sus entidades. El primero que lo haga —Colombia tiene con qué— fijará el estándar.

Riesgos.

- **Gobernanza de papel.** Políticas de datos aprobadas y archivadas, comités que no deciden, inventarios desactualizados. El antídoto es medir: pocas métricas, públicas, trimestrales.
- **Centralización sin salvaguardas.** La tentación de resolver la fragmentación con una gran base unificada; la experiencia brasileña de 2019 muestra el riesgo simétrico al de la dispersión. La arquitectura correcta comparte con registro y propósito, no acumula.
- **Dependencia de proveedores no gobernada.** Datos críticos en infraestructuras de terceros sin cláusulas de auditoría, portabilidad ni continuidad: el incidente de 2023 es el precedente, no la excepción posible.
- **Privacidad como trámite.** Tratar la Ley 1581 como formulario de autorización y no como diseño del tratamiento; la factura llega en sanciones, tutelas y pérdida de confianza.

Perspectivas.

Con corte a mayo de 2024: la entrada en vigor del Reglamento de IA europeo (esperada para agosto de 2024) activará el calendario que hará exigibles los requisitos de datos; la política nacional de IA colombiana, en elaboración, tendrá en la infraestructura de datos su habilitador crítico; y la agenda regional de intercambio de datos (espacios de datos, identidad digital) seguirá avanzando con financiamiento multilateral. La predicción de este documento es verificable: en cinco años, la diferencia entre entidades que aprovechan la IA y entidades que la padecen se explicará menos por los modelos que compren y más por el estado de sus datos.

Recomendaciones.

La ruta de 90 días.

Días	Acción	Producto verificable
1–15	Designar responsable de datos y comité	Acto administrativo publicado
15–45	Levantar el inventario de bases de datos críticas	Catálogo inicial: dueño, contenido, base legal, sistemas
30–60	Clasificar seguridad y revisar contratos de custodia	Mapa de riesgos de terceros; cláusulas mínimas
45–75	Definir estándares de identificadores y metadatos para dominios críticos	Diccionario de datos v1
60–90	Formalizar los intercambios de datos existentes	Acuerdos escritos por cada flujo activo
90	Aprobar la política de datos y las métricas de calidad	Política de una página + tablero trimestral

Tabla 2. Ruta Enpath de 90 días. Fuente: Enpath.

Tres reglas acompañan la ruta: **empezar por los dominios que deciden** (los datos que alimentan las decisiones misionales, no los más fáciles); **medir poco y en público** (tres métricas de calidad publicadas valen más que un manual de cien páginas); y **no comprar plataformas antes del inventario** (la herramienta llega cuando se sabe qué debe gestionar).

Metodología.

Revisión de fuentes públicas con corte al 31 de mayo de 2024: normativa colombiana (leyes 1266, 1581, 1712; documentos CONPES 3920 y 3975; Plan Nacional de Infraestructura de Datos; Decreto 767 de 2022), marcos internacionales (OCDE 2019; reglamentos europeos de gobernanza de datos, de datos y de IA), índices comparados (OURdata 2019 y 2023) y estudios regionales (OCDE–CAF 2022; ILIA 2023). El marco de cuatro capas es una síntesis de Enpath a partir de los estándares citados (OCDE, DAMA-DMBOK) adaptada a la escala de entidades públicas latinoamericanas.

Limitaciones. No existe medición pública de la gobernanza interna de datos por entidad en Colombia ni en la región; el diagnóstico institucional se apoya en evidencia indirecta (estudios regionales, índices, casos documentados) y en la práctica profesional. La ruta de 90 días es una propuesta normativa no validada experimentalmente; su calendario supone una entidad de tamaño medio con equipo de TI propio.

Referencias.

1. AGESIC. *Documentación institucional: interoperabilidad y gobernanza digital*. Gobierno de Uruguay. https://www.gub.uy/agencia-gobierno-electronico-sociedad-informacion-conocimiento/
2. CENIA y CEPAL (2023). *Índice Latinoamericano de Inteligencia Artificial (ILIA) 2023*. https://indicelatam.cl/
3. Congreso de la República de Colombia. *Ley 1266 de 2008; Ley 1581 de 2012; Ley 1712 de 2014*. Diario Oficial. https://www.funcionpublica.gov.co/eva/gestornormativo/norma.php?i=49981
4. Consejo de la Unión Europea (2024). *Artificial intelligence act: Council gives final green light*. 21 de mayo de 2024. https://www.consilium.europa.eu/en/press/press-releases/2024/05/21/artificial-intelligence-ai-act-council-gives-final-green-light-to-the-first-worldwide-rules-on-ai/
5. DNP (2018). *Documento CONPES 3920: Política Nacional de Explotación de Datos (Big Data)*. https://colaboracion.dnp.gov.co/CDT/Conpes/Econ%C3%B3micos/3920.pdf
6. e-Estonia. *X-Road: interoperability services* (documentación institucional). https://e-estonia.com/solutions/interoperability-services/x-road/
7. MinTIC (2021). *Plan Nacional de Infraestructura de Datos*. Gobierno de Colombia. https://www.mintic.gov.co/portal/inicio/Sala-de-prensa/Noticias/161810:Gobierno-nacional-presenta-el-Plan-Nacional-de-Infraestructura-de-Datos
8. OCDE (2019). *The Path to Becoming a Data-Driven Public Sector*. OECD Publishing. https://doi.org/10.1787/059814a3-en
9. OCDE (2023). *2023 OECD Open, Useful and Re-usable data (OURdata) Index*. Diciembre de 2023. https://www.oecd.org/en/publications/2023-oecd-open-useful-and-re-usable-data-ourdata-index_a37f51c3-en.html
10. OCDE y CAF (2022). *The Strategic and Responsible Use of Artificial Intelligence in the Public Sector of Latin America and the Caribbean*. https://doi.org/10.1787/1f334543-en
11. Unión Europea (2022). *Reglamento (UE) 2022/868 de Gobernanza de Datos*. https://eur-lex.europa.eu/eli/reg/2022/868/oj
12. Unión Europea (2024). *Reglamento (UE) 2023/2854 de Datos*. https://eur-lex.europa.eu/eli/reg/2023/2854/oj

---

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