

Ley que regula los datos sintéticos, los entornos seguros de acceso a datos públicos y la reidentificación para la investigación y la inteligencia artificial

Habilita entornos seguros de acceso para investigadores acreditados a datos públicos disociados, con excepción expresa al secreto estadístico; encarga a la ANPD un estándar de anonimización y datos sintéticos con presunción refutable; y tipifica la reidentificación como infracción muy grave, sin sancionar la investigación de seguridad de buena fe.

Comisión (Cámara de Diputados): Justicia y Derechos Humanos (+ Ciencia, Innovación Tecnológica y Sociedad Digital) ·

Revisora (Senado): Justicia y Derechos Humanos · **Prioridad:** media · **Costo fiscal:** bajo · **Plazo:** mediano

Plan Perú 2050: Inteligencia Artificial, Ciencia y Tecnología, Perú en La Era Digital

I. Exposición de motivos

La investigación y el desarrollo de IA requieren datos detallados (microdatos de salud, educación, programas sociales, registros administrativos), pero su apertura choca legítimamente con la protección de datos personales. La Ley 29733 define la anonimización (irreversible) y la disociación (reversible) y exige de consentimiento el tratamiento de datos anonimizados o disociados (art. 14, numeral 8), y el Reglamento aprobado por D.S. 016-2024-JUS encarga a la Autoridad Nacional de Protección de Datos Personales aprobar disposiciones complementarias u orientativas sobre disociación y anonimización (art. 41.3). Sin embargo, no existen reglas legales sobre datos sintéticos —que pueden memorizar y reproducir registros reales— ni sobre entornos seguros donde investigadores acreditados analicen datos disociados sin extraerlos, y el secreto estadístico del Decreto Legislativo 604 impide al INEI habilitar ese acceso sin norma de rango legal. Como consecuencia, muchas entidades optan por no compartir datos ante el riesgo de sanción, o los comparten sin garantías suficientes. Además, la reidentificación de personas a partir de datos supuestamente anónimos, técnicamente más sencilla con IA, no está tipificada como infracción imputable a cualquier persona, y una prohibición mal diseñada podría sancionar a investigadores que, de buena fe, detectan anonimizaciones defectuosas. El Plan IA Perú 2050 (CIP/CNPP) plantea los datos como activo nacional. Las guías de la ICO británica, el NIST (SP 800-188) y el Reglamento (UE) 2022/868 de gobernanza de datos ofrecen modelos verificables.

II. Fórmula legal

EL CONGRESO DE LA REPÚBLICA ha dado la Ley siguiente:

LEY QUE REGULA LOS DATOS SINTÉTICOS, LOS ENTORNOS SEGUROS DE ACCESO A DATOS PÚBLICOS Y LA REIDENTIFICACIÓN PARA LA INVESTIGACIÓN Y LA INTELIGENCIA ARTIFICIAL

Artículo 1. Objeto

La presente ley tiene por objeto habilitar la reutilización segura de datos públicos para la investigación y el desarrollo de inteligencia artificial, mediante reglas sobre datos sintéticos, entornos seguros de acceso y reidentificación, garantizando la protección de datos personales.

Artículo 2. Definiciones

2.1. Los términos anonimización y disociación tienen el significado que les atribuye el artículo 2 de la Ley 29733, Ley de Protección de Datos Personales.

2.2. Para efectos de la presente ley se entiende por: a) Datos sintéticos: datos generados artificialmente a partir de modelos estadísticos o de inteligencia artificial que reproducen propiedades de un conjunto original sin corresponder a personas reales; b) Entorno seguro de acceso: infraestructura controlada donde investigadores acreditados analizan datos disociados sin poder extraer datos individuales.

Artículo 3. Estándar de anonimización y de datos sintéticos

3.1. La Autoridad Nacional de Protección de Datos Personales, en coordinación con el Instituto Nacional de Estadística e Informática (INEI) y la Secretaría de Gobierno y Transformación Digital de la Presidencia del

Consejo de Ministros, aprueba mediante resolución publicada en el diario oficial El Peruano el estándar técnico de anonimización y de generación y evaluación de datos sintéticos, que incluye pruebas de riesgo de reidentificación y de inferencia de pertenencia y métricas de utilidad. El estándar se revisa al menos cada dos años y es el estándar nacional de referencia para la anonimización de datos públicos.

3.2. Los datos anonimizados y los datos sintéticos que superen las pruebas del estándar se presumen no personales para efectos de la Ley 29733, salvo que se acredite que permiten identificar a una persona con medios razonablemente previsibles. El responsable conserva la documentación de las pruebas y repite la evaluación cuando cambien las técnicas disponibles. Si se comprueba que un conjunto permite identificar a personas, recupera la condición de dato personal y la entidad suspende su difusión y aplica nuevamente el procedimiento.

Artículo 4. Entornos seguros de acceso

4.1. El INEI y las entidades que custodien registros administrativos de alto valor pueden implementar entornos seguros de acceso para investigadores acreditados, preferentemente mediante un entorno centralizado con acceso remoto para varias entidades.

4.2. El acceso se otorga por proyecto, con finalidad de interés público o científico, compromiso de confidencialidad, registro de consultas y control de los resultados antes de su extracción para impedir la divulgación de casos individuales. La relación de proyectos autorizados se publica.

4.3. Los entornos seguros cumplen como mínimo: autenticación multifactor; ausencia de canales de salida de datos no autorizados; registro y conservación de las consultas y accesos; cifrado en reposo; y notificación de incidentes al Centro Nacional de Seguridad Digital y a la Autoridad Nacional de Protección de Datos Personales conforme a la normativa vigente.

4.4. El acceso a datos disociados en entornos seguros de acceso conforme a la presente ley no vulnera el secreto estadístico regulado en el artículo 97 del Decreto Legislativo 604, Ley de Organización y Funciones del Instituto Nacional de Estadística e Informática, siempre que no se divulguen resultados que permitan identificar a los informantes. No comprende la información protegida por el secreto bancario ni por la reserva tributaria.

Artículo 5. Acreditación de investigadores y salvaguardas para datos sensibles

5.1. Pueden acreditarse investigadores de universidades, institutos de investigación y entidades públicas, y de empresas en proyectos colaborativos con estos, conforme a los requisitos del reglamento, que incluyen capacitación en protección de datos y la suscripción de un acuerdo de uso.

5.2. Tratándose de datos sensibles, el acceso requiere además: a) evaluación de impacto en protección de datos aprobada por la entidad custodia; b) aprobación del proyecto por un comité de ética de la investigación registrado; y c) que la finalidad no pueda alcanzarse con datos anonimizados o sintéticos.

Artículo 6. Modificación del numeral 8 del artículo 14 de la Ley 29733

Se modifica el numeral 8 del artículo 14 de la Ley 29733, Ley de Protección de Datos Personales, en los términos siguientes:

«Artículo 14. Limitaciones al consentimiento para el tratamiento de datos personales

No se requiere el consentimiento del titular de datos personales, para los efectos de su tratamiento, en los siguientes casos:

(...)

8. Cuando se hubiera aplicado un procedimiento de anonimización o disociación. El tratamiento de datos disociados de entidades públicas con fines de investigación científica o estadística de interés público, incluido el desarrollo y evaluación de sistemas de inteligencia artificial, se realiza preferentemente en un entorno seguro de acceso autorizado conforme a ley, sin adoptar decisiones sobre personas concretas; el reglamento establece los casos en que puede autorizarse su transferencia fuera de dichos entornos.

(...))»

Artículo 7. Reidentificación: prohibición, infracción y excepción de buena fe

7.1. Se prohíbe reidentificar a personas a partir de datos anonimizados, disociados o sintéticos fuera de los supuestos autorizados, así como utilizar, divulgar o comercializar datos reidentificados.

7.2. Constituye infracción muy grave, imputable a cualquier persona natural o jurídica y sancionable por la Autoridad Nacional de Protección de Datos Personales con multa conforme al artículo 39 de la Ley 29733 y a

su procedimiento sancionador, la conducta prohibida en el numeral 7.1, sin perjuicio de la responsabilidad penal que corresponda.

7.3. No constituye infracción la prueba de reidentificación realizada por la entidad responsable como prueba de robustez, ni la realizada con fines de investigación o de verificación de la seguridad, siempre que: a) no se divulguen ni utilicen los datos de personas reidentificadas; b) se comuniquen el hallazgo a la entidad responsable y a la Autoridad Nacional de Protección de Datos Personales dentro de los diez días hábiles siguientes; c) se destruyan los datos reidentificados; y d) la publicación de resultados se haga de forma agregada y una vez adoptadas las medidas correctivas o transcurridos noventa días desde la comunicación.

Artículo 8. Uso de datos sintéticos en el Estado

Las entidades públicas priorizan el uso de datos sintéticos que cumplan el estándar para pruebas de software, capacitación y entrenamiento inicial de sistemas de inteligencia artificial, cuando ello sea técnicamente viable.

Primera. Disposición Complementaria Final

Reglamentación. El Poder Ejecutivo, mediante decreto supremo refrendado por el ministro de Justicia y Derechos Humanos y el presidente del Consejo de Ministros, reglamenta la presente ley en un plazo de ciento veinte días hábiles contados a partir del día siguiente de su publicación, y adecúa los artículos 41 y 134 del Reglamento de la Ley 29733, aprobado por Decreto Supremo 016-2024-JUS.

Segunda. Disposición Complementaria Final

Financiamiento. La implementación se financia con cargo al presupuesto institucional de las entidades involucradas, sin demandar recursos adicionales al Tesoro Público.

Tercera. Disposición Complementaria Final

Vigencia. La presente ley entra en vigencia al día siguiente de su publicación en el diario oficial El Peruano. Los artículos 4 y 5 son exigibles desde la vigencia del reglamento.

III. Análisis costo-beneficio

El costo es bajo a moderado: el INEI ya administra microdatos y puede ampliar sus servicios de acceso; un entorno seguro centralizado con acceso remoto para varias entidades es más eficiente que entornos por entidad. El beneficio es destrabar el uso de datos públicos para investigación en políticas sociales, salud y educación, y para entrenar y evaluar modelos con información nacional, sin sacrificar la privacidad. La tipificación de la reidentificación disuade abusos, mientras la excepción de buena fe preserva la investigación que detecta anonimizaciones defectuosas. La presunción refutable sobre datos anonimizados y sintéticos evita dejar sin tutela a las personas si el avance técnico permite reidentificarlas. Conforme al art. 79 de la Constitución, la ley no impone gasto (los entornos son facultativos, art. 4.1); requiere opinión del MEF antes del dictamen.

Normas que modifica

- Ley 29733, Ley de Protección de Datos Personales, art. 14, numeral 8 — *modifica*
- Ley 29733, Ley de Protección de Datos Personales, art. 39 (multa aplicable a nueva infracción muy grave) — *complementa*
- Decreto Legislativo 604, art. 97 (secreto estadístico) — *complementa*
- Decreto Supremo 016-2024-JUS, arts. 41 y 134 — *complementa*

Entidades implicadas (quién hace qué)

- **ANPD** (sancionadora): Aprueba y revisa el estándar (art. 3), recibe reportes de buena fe (art. 7.3) y sanciona la reidentificación (art. 7.2).
- **INEI** (ejecutora): Coordina el estándar y puede operar el entorno seguro centralizado con control de divulgación (arts. 3.1 y 4).
- **PCM-SGTD** (consultiva): Coordina el estándar y la infraestructura (art. 3.1); co-refrenda el reglamento.
- **CNSD** (supervisora): Recibe notificaciones de incidentes en entornos seguros (art. 4.3).

• **Entidades custodias de registros administrativos (MINSA, MINEDU, MIDIS, RENIEC, entre otras)** (ejecutora): Pueden habilitar acceso en entornos seguros y aprueban evaluaciones de impacto para datos sensibles (arts. 4 y 5.2).

Sujetos obligados: Entidades que habiliten entornos seguros; investigadores acreditados; cualquier persona respecto de la prohibición de reidentificar.

Beneficiarios: Investigadores, universidades y entidades públicas que analizan datos para políticas; titulares de datos por mayor protección. No cuantificable con información pública.

Impacto económico estimado (confianza media; referencial, no oficial)

Tipo: gasto

Costo para el Estado: No cuantificable con información pública

Horizonte: 2027-2031

Financiamiento: Presupuesto institucional del INEI, la ANPD y las entidades custodias, sin demandar recursos adicionales al Tesoro Público. Requiere coordinación con el Poder Ejecutivo e informe previo del MEF antes del dictamen (art. 79; STC 00018-2023-PI/TC).

Supuestos: Los entornos seguros son facultativos (art. 4.1); su costo depende de si se implementa un entorno centralizado en el INEI (escritorio virtual sin salida de datos) y de la demanda de proyectos, sin costo unitario público verificable. El estándar se elabora con personal existente de la ANPD y el INEI.

Beneficio para el Estado: Mejores evaluaciones de políticas públicas con registros administrativos; menor riesgo de brechas por compartir datos sin garantías. Cualitativo.

Beneficio para la sociedad: Investigación con datos nacionales y protección reforzada frente a la reidentificación. Cualitativo.

Inversión privada: Proyectos colaborativos empresa-academia con acceso a datos; no cuantificable.

Costo para privados: Investigadores: acreditación, acuerdos de uso y, para datos sensibles, comité de ética; costo bajo.

Riesgos: Exfiltración desde entornos seguros; datos sintéticos que memorizan registros reales; uso abusivo de la excepción de buena fe; baja adopción si no se habilita un entorno centralizado.

Revisión especializada (asistida por IA, sobre la versión anterior)

• **Constitucional** — Con observaciones: La excepción al consentimiento para investigación limita el derecho a la autodeterminación informativa (art. 2.6), pero puede superar el test de proporcionalidad si se refuerzan las salvaguardas para datos sensibles y la presunción de anonimización es refutable. Debe armonizarse con el secreto estadístico del INEI, previsto en norma con rango de ley.

• **Técnica legislativa** — Con observaciones: La propuesta es técnicamente sólida y sí modifica la Ley 29733, pero debe armonizar el acceso a microdatos con el secreto estadístico y otros secretos legales, ubicar la excepción al consentimiento en el artículo 14 vigente y tipificar la reidentificación en la norma que contiene el régimen de infracciones, verificando además quiénes son los sujetos infractores.

• **Concordancia normativa** — Con observaciones: La novedad real son los entornos seguros de acceso y el estándar de datos sintéticos. El art. 14-A propuesto duplica en gran parte el art. 14, numeral 8 de la Ley 29733 (no se requiere consentimiento cuando se aplica anonimización o disociación) y el art. 41 del D.S. 016-2024-JUS; y usa 'seudonimización', término ajeno al sistema peruano, que la denomina 'disociación' (art. 2 Ley 29733, reversible).

• **Seguridad** — Con observaciones: Propuesta sólida y necesaria; los entornos seguros son la respuesta correcta al riesgo de reidentificación. Debe fijar requisitos técnicos mínimos de los entornos, reconocer que los datos sintéticos pueden filtrar registros reales y, sobre todo, no sancionar la investigación de robustez de buena fe que detecta anonimizaciones defectuosas.

• **Derecho tecnológico** — Con observaciones: Propuesta necesaria y técnicamente bien orientada (entornos seguros, NIST SP 800-188, Reglamento (UE) 2022/868). Debe armonizar su terminología con la Ley 29733, que ya define anonimización (irreversible) y disociación (reversible, equivalente a seudonimización), evitar la presunción absoluta de que los datos sintéticos no son personales, articular con el secreto estadístico del INEI y modificar expresamente el régimen de infracciones.

IV. Efecto de la vigencia de la norma

Modifica el numeral 8 del art. 14 de la Ley 29733 (texto del numeral por verificar contra el texto vigente). Tipifica en la propia ley una infracción muy grave sancionable conforme al art. 39 de la Ley 29733 y dispone adecuar los arts. 41 y 134 del D.S. 016-2024-JUS. Establece una excepción expresa al secreto estadístico del art. 97 del D.Leg. 604 limitada a entornos seguros, sin afectar secreto bancario ni reserva tributaria. Complementa el D.L. 1412 y el D.U. 007-2020. Coordinado con IA-04 (apertura de datos) e IA-08 (datos clínicos), que remiten a este estándar. No deroga normas.

V. Indicador de éxito al 2031

Al 2031, estándar de anonimización y datos sintéticos aprobado y al menos cinco entornos seguros de acceso operativos (INEI y cuatro registros administrativos de alto valor) con proyectos autorizados publicados.

Normas relacionadas

- Ley 29733, Ley de Protección de Datos Personales — <https://www.leyes.congreso.gob.pe/Documentos/Leyes/29733.pdf>
- Decreto Supremo 016-2024-JUS, Reglamento de la Ley 29733 — https://www3.congreso.gob.pe/Docs/DGP/DIDP/files/ds_016-2024-jus.pdf
- Decreto Legislativo 604, Ley de Organización y Funciones del INEI — <https://www.leyes.congreso.gob.pe/Documentos/DecretosLegislativos/00604.pdf>
- Decreto Legislativo 1412, Ley de Gobierno Digital — <https://busquedas.elperuano.pe/dispositivo/NL/1691026-1>

Antecedentes

No se identificaron en SPLEY proyectos del periodo 2021-2026 sobre anonimización o datos sintéticos (búsqueda de la revisión de concordancia normativa; por confirmar al dictaminar). Relacionado: PL 04708/2022-CR, que modifica la Ley 29733.

Referencias internacionales

- Unión Europea: Reglamento (UE) 2022/868 de Gobernanza de Datos (reutilización de datos protegidos del sector público en entornos seguros) — <https://eur-lex.europa.eu/eli/reg/2022/868/oj>
- Estados Unidos: NIST SP 800-188, De-Identifying Government Datasets — <https://csrc.nist.gov/pubs/sp/800/188/final>
- Reino Unido: ICO, guía sobre anonimización — <https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/data-sharing/anonymisation/>

Fuentes

- Decreto Supremo 016-2024-JUS, Reglamento de la Ley 29733 (art. 41) — https://www3.congreso.gob.pe/Docs/DGP/DIDP/files/ds_016-2024-jus.pdf
- NIST SP 800-188 — <https://csrc.nist.gov/pubs/sp/800/188/final>
- Plan Perú 2050 (CNPP/CIP), dashboard de comisiones temáticas — <https://planperu2050.pe>

Propuesta ciudadana de libre uso — no es un proyecto presentado. Versión web:
<https://unimauro.github.io/agenda-legislativa-tech/p/IA-10.html>