

# Bienes comunes de datos

*Preguntas frecuentes*

11 de mayo de 2026

Stefaan Verhulst  
Andrew J. Zahuranec  
Hannah Chafetz  
Leona Verdadero  
Jennifer Hansen

## ÍNDICE

<b>Introducción: Por qué los bienes comunes de datos son importantes ahora. . . . .</b>	<b>3</b>
<b>Preguntas. . . . .</b>	<b>4</b>
¿Qué son los bienes comunes de datos? . . . . .	4
¿De dónde surge la idea de los bienes comunes? . . . . .	5
¿Cómo se gestionan los bienes comunes? . . . . .	6
¿Todos los bienes comunes de datos encarnan plenamente los ocho principios de Elinor Ostrom? . . . . .	7
¿Qué define a un bien común de datos, a diferencia de un bien común? . . . . .	7
¿En qué se diferencian los «bienes comunes de datos» de las cooperativas de datos (data collaboratives), los fideicomisos, los espacios y otras estructuras? . . . . .	8
¿En qué se diferencia un bien común de datos de los datos abiertos? . . . . .	9
¿Qué problemas pretenden resolver los bienes comunes de datos? . . . . .	10
¿Cuáles son algunas de las críticas a los bienes comunes de datos? . . . . .	10
¿Por qué es fundamental una licencia social para los bienes comunes de datos? . . . . .	13
¿Cuál es el papel de la custodia de datos en los bienes comunes de datos? . . . . .	13
¿Existen ya los bienes comunes de datos? . . . . .	14
¿Cuáles son los modelos de financiación utilizados para apoyar los bienes comunes de datos? . . . . .	15
¿Por qué es importante para la inteligencia artificial? . . . . .	16
¿Por qué es este modelo adecuado para los grupos marginados? . . . . .	17
<b>Lecturas adicionales . . . . .</b>	<b>18</b>



## INTRODUCCIÓN

# Por qué los bienes comunes de datos son importantes ahora

Una de las grandes paradojas de nuestra era datificada es que vivimos en medio de una abundancia y una escasez sin precedentes. A pesar de que los datos cobran cada vez más importancia para nuestra capacidad de promover el bien público, también siguen siendo profundamente —y quizás cada vez más— inaccesibles y controlados de forma privada. En respuesta a ello, han surgido cada vez más voces a favor de los bienes comunes de datos: fondos de datos que serían (auto)gestionados por comunidades o entidades específicas que operan en interés del público. Estos fondos podrían entonces hacerse accesibles y reutilizarse para el bien común.

*Los bienes comunes de datos ofrecen una alternativa* al uso creciente de silos de datos privatizados o a la reutilización extractiva de conjuntos de datos abiertos. Son una forma de organizar los datos como un [recurso compartido](#), gobernado colectivamente y utilizado para promover el valor público. Aunque la práctica aún está en fase incipiente y la evidencia sigue siendo limitada, los bienes comunes de datos podrían actuar como una infraestructura útil para la IA de interés público al:

- ▶ Permitir un acceso equitativo a datos de alta calidad
- ▶ Apoyar sistemas de IA basados en información diversa y representativa
- ▶ Crear vías institucionales para la reutilización responsable de los datos
- ▶ Fomentar la confianza a través de la participación, la transparencia y la rendición de cuentas

Los bienes comunes de datos no surgen de forma orgánica. Requieren una gobernanza intencional, una custodia responsable de los datos y una licencia social. Esta sección de preguntas frecuentes ofrece una visión general práctica de qué son los bienes comunes de datos, por qué son importantes y cómo diseñarlos de forma responsable.



# Preguntas

## ¿Qué son los bienes comunes de datos?

[Los bienes comunes de datos](#) son ecosistemas de datos gobernados de forma colaborativa, diseñados para reunir y proporcionar un acceso responsable (y regulado) a conjuntos de datos diversos y de alta calidad procedentes de uno o varios sectores, con el fin de permitir el desarrollo y la implementación de aplicaciones de Inteligencia Artificial (IA) generativa que aborden retos de interés público.

Este enfoque se distingue por tratar los datos como un recurso compartido gestionado por una comunidad, en lugar de controlarlos exclusivamente como propiedad individual o institucional. Además, reconoce que los datos abiertos —abiertos por defecto y, a menudo, totalmente públicos— no ofrecen, por sí solos, una solución completa. En un panorama de datos en constante evolución, garantizar que la apertura sirva a los objetivos de interés público requiere abordar cómo fluyen el reconocimiento y los beneficios, en particular cuando los datos disponibles públicamente pueden reutilizarse de forma generalizada sin que un beneficio claro revierta a las comunidades que los custodian.

En este modelo reinventado de bienes comunes de datos, comunidades específicas que operan en interés público definen las condiciones en las que se recopilan, comparten y (re)utilizan los datos. El acceso es compartido, pero solo bajo las condiciones específicas establecidas a través de estructuras de gobernanza que reflejan las prioridades y limitaciones colectivas. El uso suele estar vinculado a un fin concreto, en consonancia con las necesidades y expectativas de las comunidades contribuyentes. Muchos bienes comunes de datos también incorporan mecanismos de distribución de beneficios para garantizar que quienes proporcionan acceso al conocimiento o a los materiales participen en los beneficios y que la participación no refuerce las desigualdades existentes.

En esencia, los bienes comunes de datos no son solo un conjunto de datos o una plataforma de datos. Se trata de mecanismos institucionales que definen claramente quién tiene

autoridad sobre los datos, quién puede acceder a ellos, en qué condiciones, con qué fines, con qué garantías y supervisión, y cuáles son los resultados esperados. Esto incluye normas claras sobre el acceso, mecanismos de supervisión y cumplimiento, y estructuras de participación que van más allá de la aportación de datos para incluir la gobernanza. Su objetivo es cambiar el enfoque de los datos como recurso extractivo hacia los datos como un activo custodiado colectivamente para el bien público.

## ¿De dónde surge la idea de los bienes comunes?

Un [commons](#) o bien común es un recurso compartido que pertenece colectivamente a una comunidad o grupo de usuarios, regido por normas que dictan cómo debe utilizarse. Los bienes comunes [suelen](#) encarnar principios de autonomía local, responsabilidad mutua y beneficios compartidos. Entre los ejemplos típicos de commons se encuentran los bosques, las pesquerías y las aguas subterráneas.

El concepto de «commons» existe desde hace siglos, pero se popularizó en la era moderna con la publicación de [La tragedia de los comunes](#), de Garrett Hardin. Partiendo de los argumentos del economista del siglo XIX William Forester Lloyd, Hardin sostenía que los individuos que actúan en su propio interés se comportan en contra del bien común, y que, motivados por el interés propio, quedan atrapados en un comportamiento que perjudica el bien colectivo.

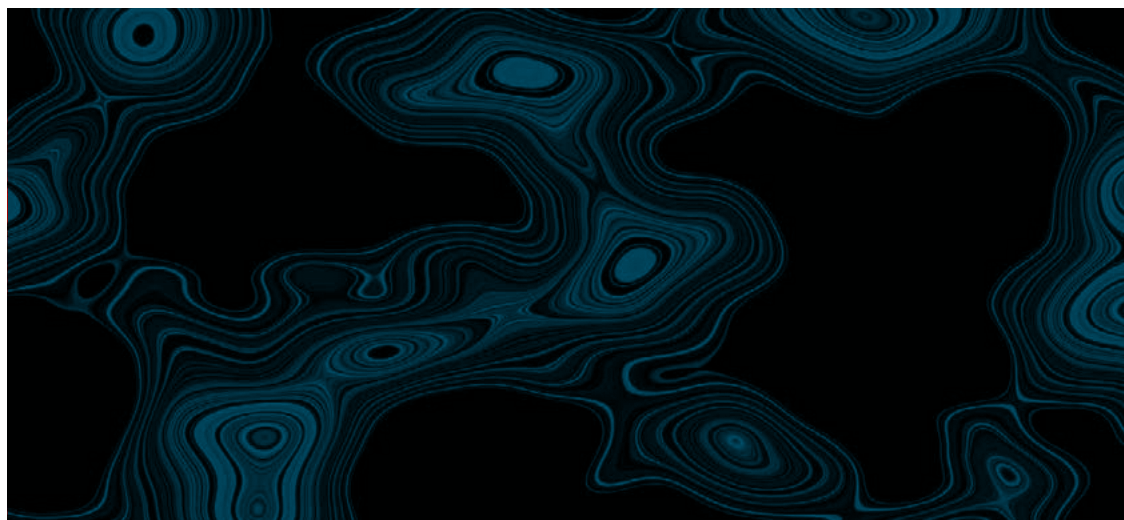
El trabajo de Lloyd y Hardin, centrado especialmente en la superpoblación y los recursos limitados, fue ampliamente refutado por Elinor Ostrom, economista y politóloga ganadora del Premio Nobel. Su trabajo, en particular su libro [Governing the Commons](#) (en español: El gobierno de los commons), demuestra que los seres humanos no están atrapados ni son impotentes, que la «tragedia» en la gestión de los recursos compartidos no es inevitable, sino que puede evitarse mediante la cooperación, la supervisión y el cumplimiento de las normas.

Aunque Ostrom ya había escrito sobre el [conocimiento](#) como [recurso común](#) a principios de la década de 2000, el concepto de «bienes comunes de datos» no surgió hasta la década de 2010. En el ámbito biomédico y de la investigación, los bienes comunes de datos se consideraban una forma de [controlar](#) los datos académicos. En 2019, la directora de la Biblioteca Nacional de Medicina, Patti Brennan, [describió](#) cómo los bienes comunes de datos estaban surgiendo como una forma importante de compartir datos, así como un «conjunto de principios, estrategias de gobernanza y herramientas». Un intento reciente de aplicar este concepto a los datos lo ha llevado a cabo la [Fundación Mozilla](#).

## ¿Cómo se gestionan los bienes comunes?

Los bienes comunes pueden adoptar muchas formas y ser utilizados por las comunidades para gestionar muchos tipos diferentes de recursos compartidos. Nuestra investigación sigue los [ocho principios de Elinor Ostrom para la gestión de los commons](#). Estos son:

- ▶ **Límites claramente definidos:** La identidad del grupo y los límites del recurso compartido están claramente definidos.
- ▶ **Congruencia entre las necesidades y las condiciones locales:** Los miembros de los bienes comunes negocian un sistema que recompensa a los miembros por sus contribuciones. Los beneficios desproporcionados deben ganarse. La desigualdad injusta envenena el esfuerzo.
- ▶ **Mecanismos de toma de decisiones colectivas:** Los miembros del grupo establecen las normas que rigen los bienes comunes y toman decisiones por consenso. Aunque haya desacuerdos individuales, todos están alineados con los objetivos generales del grupo.
- ▶ **Derechos reconocidos de organización:** Los grupos tienen autoridad propia para gestionar y supervisar sus asuntos. No están limitados por normas impuestas desde fuera que no pueden adaptarse a las circunstancias locales.
- ▶ **Supervisión eficaz:** Para evitar el parasitismo o la explotación activa, los bienes comunes supervisan las actividades para garantizar que todos los miembros cumplan las normas acordadas. Lo ideal es que esta supervisión sea de bajo coste y permita identificar rápidamente las infracciones de las normas.
- ▶ **Sanciones graduadas:** Las infracciones de las normas acarrearán consecuencias cada vez más severas.
- ▶ **Resolución accesible de conflictos:** El grupo ha acordado mecanismos que permiten resolver las disputas rápidamente y de formas que el grupo, en general, percibe como justas.
- ▶ **Entidades anidadas estructuradas:** La responsabilidad de gestionar el recurso común se ha integrado en niveles anidados. La gestión de los recursos se lleva a cabo, por defecto, en el nivel más bajo posible, pero puede vincularse a un sistema interconectado más amplio.



## ¿Todos los bienes comunes de datos encarnan plenamente los ocho principios de Elinor Ostrom?

Los ocho principios de Ostrom pueden aplicarse a diversos sectores. Sin embargo, a menudo existen como un ideal. En la práctica, muchas iniciativas no logran cumplir plenamente cada uno de estos estándares, especialmente en lo que respecta a la toma de decisiones colectiva, la autoridad de la comunidad y los mecanismos de cumplimiento.

Varios bienes comunes diseñados para datos funcionan como infraestructuras de intercambio de datos o plataformas colaborativas mantenidas por gobiernos y otras instituciones que afirman actuar en nombre del público y no son autogobernadas tal y como lo entienden los estudiosos. Pueden permitir la participación en la aportación de datos y proporcionar un acceso estructurado, pero muchos ejemplos existentes no llegan a otorgar plenamente el control a las comunidades.

## ¿Qué define a un bien común de datos, a diferencia de un bien común?

Los commons de datos presentan el siguiente conjunto de características comúnmente implementadas:

**Finalidad pública:** tiene como objetivo suministrar los datos necesarios para resolver problemas públicos y desarrollar IA de interés público.

**Gobernanza participativa:** cuenta con una estructura de gobernanza que ofrece vías significativas para que los miembros del data commons ejerzan su capacidad de decisión sobre cómo se utilizan sus datos.

**Mecanismos de rendición de cuentas:** supervisa la aplicación de las normas del data commons de forma continua. Cuenta con sistemas para gestionar los conflictos y exigir responsabilidades a quienes no cumplan las normas.

**Beneficios para los colaboradores:** ofrece beneficios claros a las partes interesadas que han contribuido al data commons. Esto puede incluir beneficios económicos, acceso anticipado a herramientas de IA o a la investigación desarrollada, atribución, información para o sobre ellos, afiliación gratuita, acceso a la infraestructura para desarrollar nuevas herramientas de IA, etc.

**Controles de acceso:** incluye normas claramente definidas que regulan cómo, cuándo y quién puede acceder a los conjuntos de datos y utilizarlos. Esto puede implicar marcos estandarizados, como las licencias Creative Commons, así como modelos de acceso por niveles, restringidos o basados en el consentimiento.

**Financiación y apoyo:** los data commons suelen financiarse como un bien público, a menudo con el apoyo del gobierno, de entidades filantrópicas o a través de un modelo de afiliación o cooperativo.

## ¿En qué se diferencian los «bienes comunes de datos» de las cooperativas de datos (data collaboratives), los fideicomisos, los espacios y otras estructuras?

El panorama actual de los datos ha sido testigo de [una profusión de modelos de gobernanza](#). Estos conceptos suelen tratarse como alternativas que compiten entre sí, caminos diferentes hacia un mismo objetivo. De hecho, este enfoque oculta la realidad de que cada modelo aborda retos de gobernanza distintos e incluye diferentes modelos de autoridad, participación, distribución del valor y riesgo. Siete enfoques han cobrado especial relevancia, cada uno de los cuales aborda un reto de gobernanza concreto al que se enfrentan las asociaciones para la reutilización de datos:

- ▶ **Los intermediarios de datos**, que actúan como agentes para reducir los costes de transacción y facilitar el intercambio de confianza entre los titulares de datos y los usuarios. Son especialmente útiles para abordar los **fallos de coordinación**, ya que ofrecen servicios de intermediación y reducción de los costes de transacción mediante normas negociadas y la intermediación operativa.
- ▶ **Uniones o coaliciones de datos**, que agrupan a los participantes para reforzar el poder de negociación colectiva en los mercados de datos. Son útiles para resolver las **asimetrías de negociación** reequilibrando el poder y articulando las condiciones de uso de los datos.
- ▶ **Los fideicomisos de datos**, que delegan la autoridad de toma de decisiones en fiduciarios independientes que actúan en nombre de beneficiarios definidos. Estos fideicomisos abordan los **déficits de legitimidad** mediante la creación de una administración fiduciaria independiente que garantiza la neutralidad y compromisos creíbles.
- ▶ **Cooperativas de datos (data collaboratives)**, que integran la propiedad democrática y el control de los miembros directamente en los acuerdos de gobernanza de datos. Las cooperativas abordan la **desigualdad en la propiedad** al permitir la propiedad democrática y otorgar a los miembros la capacidad de redistribuir el control y el valor.
- ▶ **Los entornos de pruebas de datos**, que crean entornos experimentales controlados para poner a prueba las normas de gobernanza y los usos de los datos antes de su ampliación. Los entornos de pruebas de datos ayudan a abordar la **incertidumbre del sistema** al ofrecer un entorno para poner a prueba la gobernanza, las normas, las salvaguardias y la viabilidad institucional.
- ▶ **Espacios de datos**, que proporcionan una infraestructura federada y estándares compartidos para permitir la interoperabilidad entre sectores o jurisdicciones. Los espacios de datos ayudan a abordar la complejidad de la ampliación al ofrecer sistemas para la coordinación federada y fomentar la interoperabilidad entre actores y jurisdicciones.

- ▶ Los bienes comunes de datos, que establecen estructuras de gobernanza participativa para la gestión y custodia compartida de conjuntos de datos agrupados. Los commons de datos abordan las necesidades de gobernanza colectiva al ofrecer entornos para la elaboración participativa de normas y la gestión y custodia compartida en beneficio colectivo.

Un bien común de datos puede coexistir con uno de estos otros modelos o surgir de él (por ejemplo, un bien común de datos puede evolucionar a partir de un entorno de pruebas, ser facilitado por un intermediario o ampliarse hasta convertirse en un espacio de datos), pero es distinto de ellos.

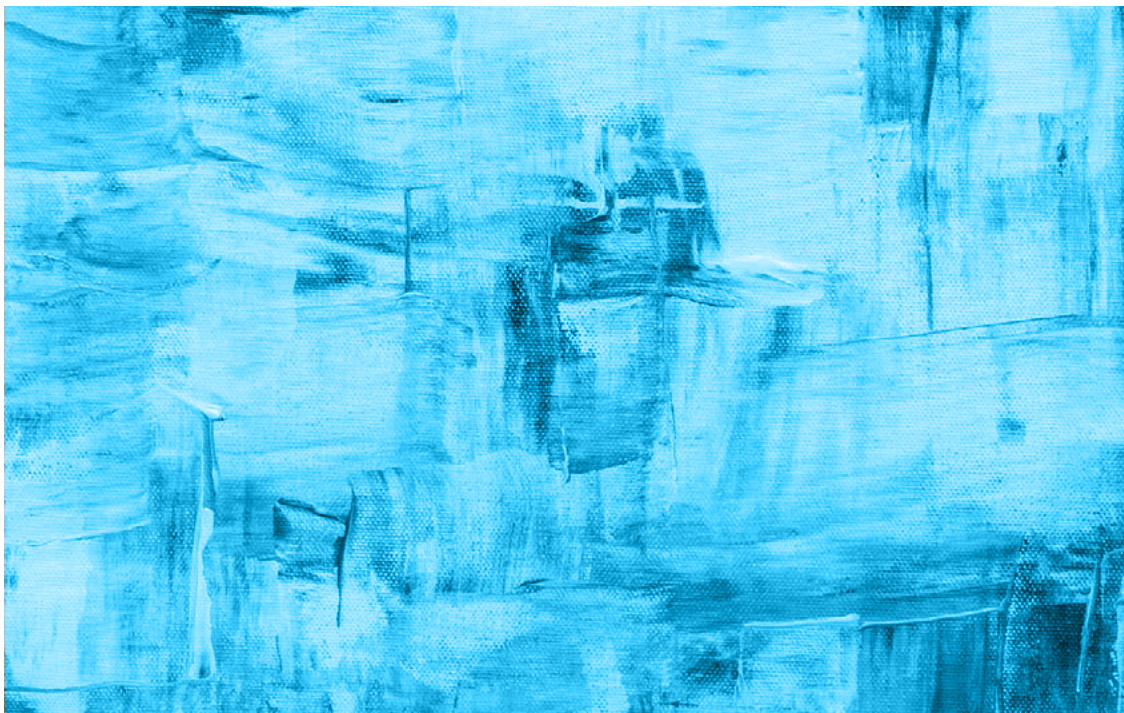
### ¿En qué se diferencia un bien común de datos de los datos abiertos?

El modelo de bienes comunes de datos tiene como objetivo abordar el reto de las asimetrías de acceso (detalladas más adelante), en las que las organizaciones hacen que sus datos sean ampliamente accesibles, lo que da lugar a que otros los recopilen y extraigan sin obtener ningún beneficio a cambio.

Si bien los bienes comunes *pueden* hacer que los datos sean abiertos para uso general, su estructura permite a sus organizadores determinar qué datos ponen a disposición, con qué finalidad y en qué condiciones. En lugar de exigir un modelo de «abierto por defecto», los bienes comunes de datos combinan la accesibilidad con la gobernanza.

Ofrecen formas estructuradas de determinar y proporcionar acceso a los datos —públicos, privados o híbridos— al tiempo que especifican las condiciones de uso, las obligaciones de reparto de beneficios, los mecanismos de supervisión y los sistemas de atribución y compensación.

Pretenden replantear la apertura en una era en la que el contexto, la licencia social y el control importan tanto como la disponibilidad.



## ¿Qué problemas pretenden resolver los bienes comunes de datos?

Los bienes comunes de datos pretenden abordar varios fallos estructurales:

**Asimetrías de acceso:** Muchas grandes organizaciones buscan utilizar datos abiertos para extraer valor de otros, al tiempo que limitan el acceso a sus propios datos privados. Los bienes comunes de datos pueden introducir mecanismos a través de los cuales las comunidades puedan influir en quién accede a sus datos, cómo y con qué propósito.

**Fallos de coordinación:** A menudo existen altos costes de transacción para compartir datos con un único actor o de forma puntual (por ejemplo, la formulación de acuerdos o normas de intercambio de datos a medida). Los bienes comunes de datos pueden crear procesos más estandarizados para compartir datos bajo normas coherentes.

**Déficit de confianza:** En muchos acuerdos de intercambio de datos, los intereses y expectativas del público o de las comunidades contribuyentes no se incorporan de manera significativa. Esto erosiona la licencia social y limita la legitimidad de cualquier (re)uso de datos. Los bienes comunes de datos tratan de dar prioridad a estas preocupaciones a través de la gobernanza, la supervisión y las condiciones de acceso.

**Infrautilización:** un problema que afecta a la reutilización de datos es que los conjuntos de datos más valiosos no se utilizan para el bien público, sino que se restringen. Los bienes comunes de datos pueden ampliar el acceso a un conjunto más amplio de actores, siempre que estos se comprometan a respetar las normas establecidas por los organizadores.

De este modo, los bienes comunes de datos pueden apuntar hacia relaciones de intercambio de datos más sostenibles y equitativas en pro del bien público. Pueden abordar algunas de las preocupaciones que alimentan la escasez de datos y el actual «invierno de los datos». Pueden reorientar el debate sobre la IA para centrarse en la IA de interés público y la toma de decisiones.

10

## ¿Cuáles son algunas de las críticas a los bienes comunes de datos?

Las críticas a los bienes comunes de datos son importantes y necesarias. Como sostienen [Tommaso Fia y Gijs van Maanen](#), entre otros, ciertos debates en torno a los bienes comunes de datos invocan el lenguaje de la «comunidad» sin especificar adecuadamente qué significa realmente «comunidad», quién la define, quién se beneficia y quién tiene la capacidad de participar. También advierten contra el «commonswashing» y el «community-washing», en los que la retórica de la gobernanza colectiva se utiliza principalmente con fines de reputación, sin redistribuir de manera significativa el poder ni los beneficios.

Estas críticas deben tomarse en serio. De hecho, existe el riesgo de que algunas iniciativas de bienes comunes de datos se vuelvan excesivamente tecnocráticas, abstractas o «solucionistas», tratando la gobernanza como un problema de arquitectura técnica e ignorando las

asimetrías políticas, institucionales y económicas más profundas. También existe una tendencia en algunos debates sobre políticas públicas a elaborar largos inventarios de modelos de gobernanza (véase más arriba) —fideicomisos de datos, cooperativas de datos, bienes comunes de datos, espacios de datos, intermediarios— como si fueran opciones intercambiables de un menú, sin cuestionar lo suficiente los valores, las condiciones institucionales o las estructuras de poder que los sustentan.

Al mismo tiempo, existe una tendencia creciente entre algunos críticos a adoptar una especie de purismo de la gobernanza: rechazar por completo la experimentación con la innovación institucional a menos que ya existan condiciones perfectas de equidad, representación o legitimidad democrática. Si bien las preocupaciones que motivan esta cautela son comprensibles, tal enfoque corre el riesgo de producir parálisis precisamente en un momento en que los sistemas de mercado y regulatorios existentes están demostrando su incapacidad para gobernar los datos y la IA de manera que se generen beneficios sociales compartidos.

La realidad es que muchos de los perjuicios actuales asociados al desarrollo extractivo de la IA y al acceso asimétrico a los datos son, en sí mismos, el resultado de una innovación insuficiente en la gobernanza institucional. Limitarse a confiar en las estructuras reguladoras existentes o en los incentivos de mercado no será suficiente para abordar los retos relacionados con el acceso a los datos de interés público, el reparto de beneficios, la gobernanza participativa o la agencia colectiva en la era de la IA.

Por lo tanto, los bienes comunes de datos no deben entenderse como una solución milagrosa ni como un sustituto de la regulación. Tampoco deben idealizarse como algo intrínsecamente democrático o equitativo. Más bien, deben considerarse un instrumento de gobernanza más dentro de un conjunto más amplio de herramientas institucionales, junto con la supervisión pública, los mecanismos fiduciarios, la gobernanza participativa y las infraestructuras públicas.

Lo que importa no es si una iniciativa se autodenomina «banco de bienes comunes de datos», sino si promueve de manera significativa:

- ▶ la agencia colectiva en materia de gobernanza de datos;
- ▶ una distribución más equitativa de los beneficios;
- ▶ la transparencia y la rendición de cuentas;
- ▶ la toma de decisiones participativa;
- ▶ la gestión y custodia sostenible de los recursos compartidos; y
- ▶ mecanismos para contrarrestar las dinámicas extractivas.

En la práctica, esto significa ser extremadamente concreto sobre varias cuestiones:

- ▶ ¿Existe realmente la comunidad más allá de la retórica?
- ▶ ¿Quién define las normas de pertenencia y gobernanza?
- ▶ ¿De qué recursos y capacidades disponen los participantes para gobernar de manera eficaz?
- ▶ ¿Existen mecanismos para abordar las asimetrías existentes de poder y conocimientos especializados?
- ▶ ¿Quién se beneficia de la colaboración?
- ▶ ¿Qué formas de rendición de cuentas, impugnación y salida existen?
- ▶ ¿Mejoran los resultados de manera significativa la equidad, el acceso o el valor público?

Es importante destacar que las respuestas a estas preguntas variarán en función del contexto. Un conjunto de bienes comunes de datos construido en torno a la soberanía de los datos indígenas, por ejemplo, plantea consideraciones de gobernanza fundamentalmente diferentes a las de un conjunto de bienes comunes de datos sobre movilidad para la planificación urbana o una colaboración de datos sanitarios para la investigación. Como subrayan Fia y van Maanen, las comunidades difieren en sus historias, identidades, vínculos sociales, capacidades de gobernanza y aspiraciones políticas.

El objetivo, por lo tanto, no es imponer un modelo de gobernanza universal, sino desarrollar mecanismos de gobernanza que sean sensibles al contexto, estén institucionalmente arraigados y sean capaces de generar beneficios compartidos legítimos. En muchos casos, esto requerirá combinar enfoques basados en los commons con una regulación más estricta, instituciones públicas, funciones de gobernanza de datos y mecanismos participativos.

En última instancia, la crítica a los bienes comunes de datos no debería llevarnos a abandonar la experimentación con la gobernanza colectiva. Debería impulsarnos a diseñar estos sistemas de forma más rigurosa, más transparente y más honesta políticamente. La alternativa —dejar la gobernanza de los datos enteramente en manos de los mercados o de la autoridad estatal centralizada— ya ha demostrado sus limitaciones.



## **¿Por qué es fundamental una licencia social para los bienes comunes de datos?**

Una [licencia social](#) es una aceptación social implícita de una actividad basada en la confianza, la legitimidad y la percepción de alineación con los valores de la comunidad. Se diferencia de las licencias de datos abiertos, ya que no es de naturaleza jurídica, sino más bien un acuerdo o plan que se desarrolla a través de la participación y la colaboración de las partes interesadas. En el contexto de los datos, capta el consentimiento colectivo centrándose en el respaldo tácito de las personas y las comunidades para la recopilación, el uso y la reutilización de sus datos.

El concepto de licencia social reconoce que el cumplimiento legal y el consentimiento son necesarios, pero no suficientes. Mientras que los regímenes legales existentes buscan el acuerdo individual en el momento de la recopilación de datos, una licencia social busca crear un proceso para que las comunidades den forma y supervisen la reutilización continua de los datos.

Una licencia social garantiza que los datos reflejen los intereses colectivos. Esto, a su vez, puede hacer que el uso de los datos se perciba como legítimo y permita procesos que generen una rendición de cuentas continua. En lugar de un consentimiento puntual —un simple proceso de aceptación o rechazo—, la licencia social anima a los actores a pensar a largo plazo, a revisar periódicamente las necesidades y expectativas para garantizar que todas las actividades sigan contando con la aprobación general de las partes interesadas de la comunidad.

Una licencia social puede ser fundamental cuando se trata de importantes disparidades de poder entre los actores, disparidades que crean vulnerabilidad o riesgo de explotación. También puede ser útil para orientar los esfuerzos no hacia necesidades a corto plazo (por ejemplo, ¿aprueba este actor este trabajo en este momento?), sino hacia intereses a largo plazo (por ejemplo, ¿qué se necesita para garantizar una participación continuada?).

13

## **¿Cuál es el papel de la custodia de datos en los bienes comunes de datos?**

Los custodios de datos (data stewards) son líderes responsables en materia de datos, facultados por sus organizaciones para crear valor público a través de intercambios de datos intersectoriales. Los custodios de datos actúan como la columna vertebral operativa de los bienes comunes de datos, ayudando a facilitar e institucionalizar el acuerdo.

Por lo general, los custodios de datos se encargarían de facilitar el acceso y negociar las condiciones en las que se podrían intercambiar los datos. También garantizarían el cumplimiento normativo y el uso ético, ayudarían a comunicar los intereses de los diferentes proveedores y usuarios de datos y trabajarían para mantener la confianza entre las partes a lo largo del tiempo.

## ¿Existen ya los bienes comunes de datos?

Los bienes comunes de datos son una práctica emergente. Son relativamente pocas las iniciativas existentes que se ajustan plenamente a la definición de sistemas gobernados de forma colaborativa en los que las comunidades ejercen autoridad sobre los datos. En cambio, los ejemplos actuales abarcan un amplio espectro de modelos que incorporan algunas características propias de los bienes comunes, como el acceso compartido, la participación de múltiples partes interesadas o la orientación hacia el interés público, sin que siempre se aplique una gobernanza colectiva.

Las iniciativas que se indican a continuación ilustran esta variedad:

**CLARIN:** CLARIN proporciona la infraestructura para combinar datos lingüísticos de diversas instituciones europeas con fines de investigación. La gobernanza es multipartita a nivel institucional, pero la autoridad en la toma de decisiones la ejercen los Estados miembros y las organizaciones, en lugar de una comunidad contribuyente más amplia.

**Common Voice:** Common Voice es un conjunto de datos de voz abierto y de colaboración colectiva que puede utilizarse para entrenar aplicaciones de voz impulsadas por IA. La iniciativa tiene como objetivo ampliar el acceso a los datos de voz para idiomas distintos del inglés y otros grupos que suelen estar infrarrepresentados en los conjuntos de datos de voz. La iniciativa está liderada por la Fundación Mozilla.

**Lacuna Fund:** El Lacuna Fund impulsa el desarrollo de conjuntos de datos etiquetados sobre países de ingresos bajos y medios que puedan utilizarse para el aprendizaje automático. El equipo acepta propuestas para la creación de conjuntos de datos en cuatro ámbitos: agricultura, idiomas, salud y datos climáticos.

**Language Data Commons of Australia:** Language Data Commons of Australia es una colaboración entre Australian Research Data Commons y la Facultad de Lenguas y Culturas de la Universidad de Queensland, junto con otros socios (por ejemplo, [First Languages Australia](#)), que busca poner a disposición datos lingüísticos australianos para «usos tanto académicos como no académicos».

**Querido Diário:** Open Knowledge Brasil pretende desarrollar un data commons para los boletines oficiales municipales. Su plataforma Querido Diário tiene como objetivo facilitar el acceso a estos conjuntos de datos y ayudar a analizar la información que contienen. Su objetivo es mejorar la toma de decisiones a nivel local en Brasil.



**Malawi Voice Data Commons:** El Malawi Voice Data Commons del Programa de Investigación y Educación para la Paz de la Universidad de Nueva York (NYU), desarrollado en colaboración con Ushahidi, el PNUD y la Fundación Mozilla, permite a los habitantes de las zonas rurales de Malawi informar de emergencias en sus lenguas nativas, creando conjuntos de datos multilingües y preparados para la inteligencia artificial con fines de respuesta humanitaria y preservación lingüística. La fase piloto se llevará a cabo en Malawi, con planes de expandirse por toda el África subsahariana.



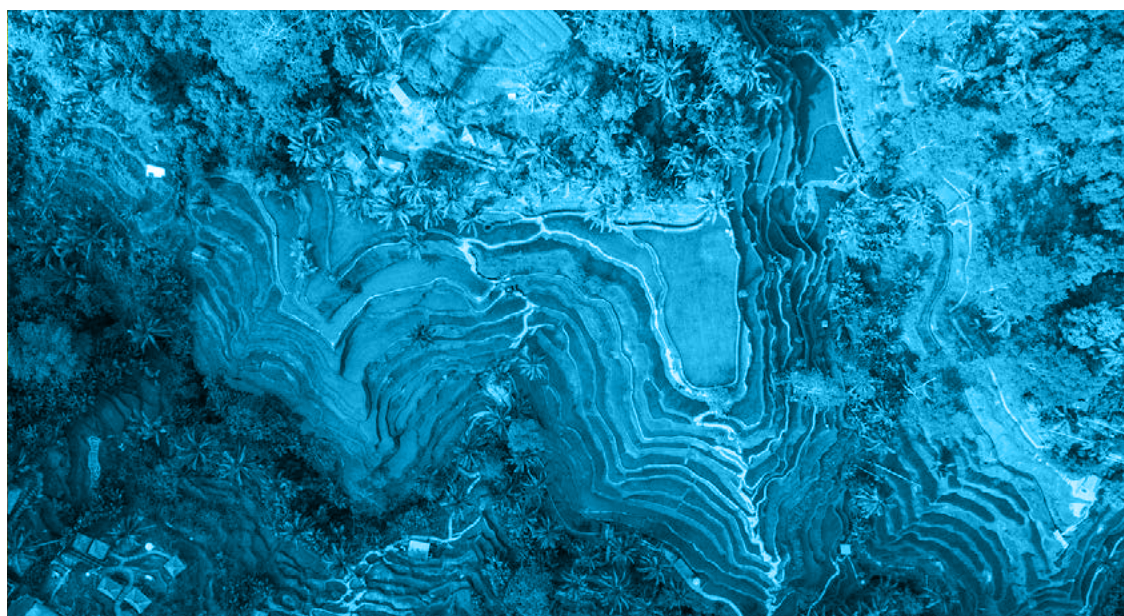
## ¿Cuáles son los modelos de financiación utilizados para apoyar los bienes comunes de datos?

La investigación en este ámbito aún está en sus inicios. Sin embargo, hemos identificado algunas estructuras de financiación recurrentes:

**Modelos liderados por el gobierno:** Los bienes comunes de datos cuentan con el apoyo de organismos gubernamentales, normalmente en colaboración con otras entidades. Estos modelos se benefician de compromisos de financiación a largo plazo.

**Modelos filantrópicos:** Los bienes comunes de datos cuentan con el apoyo de donantes del sector privado y de la sociedad civil. Estos bienes comunes combinan el capital de los donantes con la participación de la comunidad, lo que se ajusta al modelo de beneficio indirecto identificado en la literatura: el valor reside en el cumplimiento de la misión y en la implicación del ecosistema, más que en la monetización.

**Modelos de afiliación y cooperativos:** Los usuarios o colaboradores (incluidos los gobiernos) se convierten en accionistas o miembros, pasando a ser, en la práctica, “copropietarios” de los bienes comunes. Estos modelos internalizan la gobernanza y la sostenibilidad a través de la propiedad colectiva, haciendo eco de los arquetipos empresariales «freemium» y «razor-blade», en los que el acceso es abierto, pero los servicios, herramientas o privilegios de gobernanza adicionales requieren una contribución o el pago de cuotas.



## ¿Por qué es importante para la inteligencia artificial?

Los bienes comunes de datos pueden ser valiosos en la era de la inteligencia artificial. Ofrecen una infraestructura fundamental para el acceso y la reutilización de datos y pueden garantizar que este uso se ajuste a los intereses de la comunidad. Describimos esto en nuestro [Plan de Acción \(Blueprint\)](#), que señala que los commons de datos pueden ser útiles para el desarrollo de la IA, ya que pueden:

- ▶ Reunir conjuntos de datos dispares y variados necesarios para el desarrollo y la implementación de la IA;
- ▶ Proporcionar la infraestructura necesaria para estandarizar los datos en formatos preparados para la IA;
- ▶ Reducir los costes de estandarización y ayudar a evitar la duplicación de esfuerzos;
- ▶ Aumentar la diversidad cultural de los conjuntos de datos disponibles;
- ▶ Poner en práctica nuevos modelos de reutilización de datos para el bien común;
- ▶ Establecer mecanismos de procedencia de datos que mejoren la transparencia y la trazabilidad dentro de los sistemas de IA; y
- ▶ Ayudar a las comunidades de desarrolladores a crear aplicaciones de IA que respondan a necesidades reales.

16



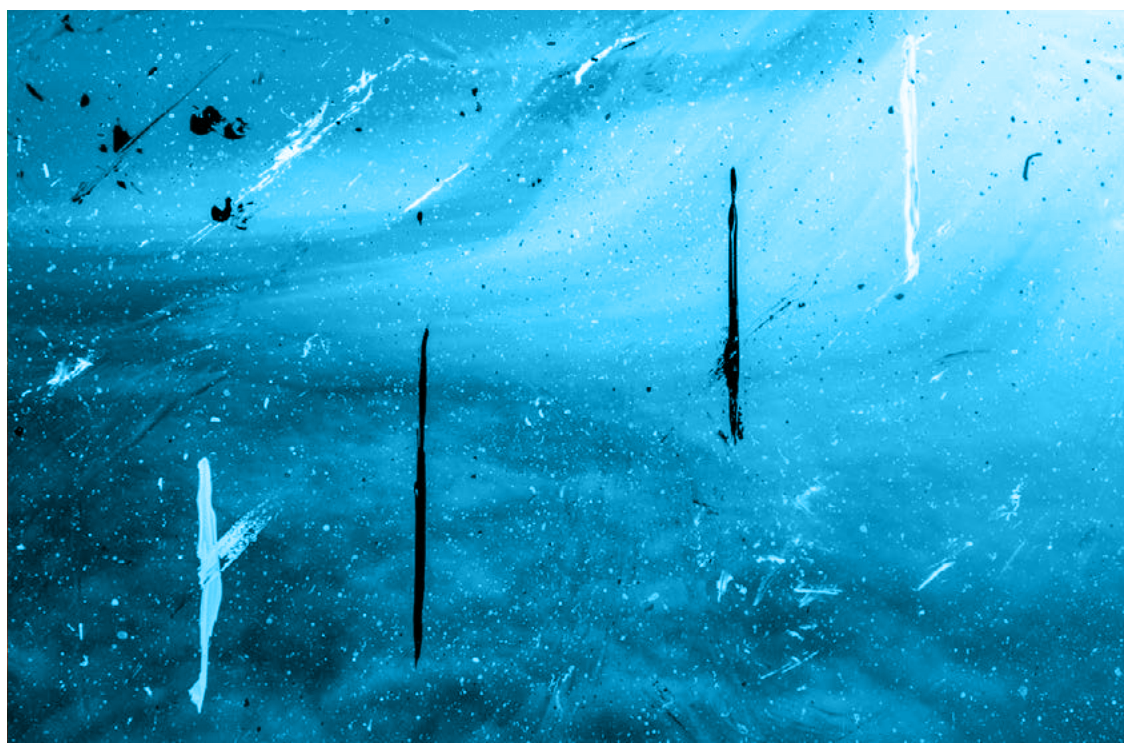
## ¿Por qué es este modelo adecuado para los grupos marginados?

Un problema importante con el desarrollo de la IA y otras herramientas basadas en datos es que pueden ser extractivas. Los datos que se ponen a disposición de una comunidad se recopilan a gran escala para entrenar sistemas propietarios, con escaso interés en reconocer o compensar a esas comunidades. El resultado es que unas pocas grandes organizaciones obtienen más activos, lo que aumenta las asimetrías de poder, y las personas tienen poco control, ya que la información sobre ellas se utiliza de formas que socavan sus intereses o van en contra de sus deseos explícitos.

A este problema lo denominamos «*instrumentalización de la apertura como arma*». Es un problema con el que están familiarizadas muchas comunidades indígenas que han visto cómo las empresas farmacéuticas extraían sus conocimientos tradicionales y su biodiversidad sin reconocimiento ni contraprestación alguna.

Los bienes comunes de datos *commons* ofrecen un marco para combinar la accesibilidad con la gobernanza. Pueden proporcionar formas estructuradas de acceder a los datos —ya sean públicos, privados o híbridos— al tiempo que especifican las condiciones de uso, las obligaciones de distribución de beneficios, los mecanismos de supervisión y los sistemas de atribución y compensación. Esto puede garantizar que los datos se utilicen de formas que se ajusten a las expectativas, necesidades e intereses de una comunidad.

Sin embargo, no es una panacea. Debe existir en un entorno en el que la comunidad esté bien definida, cuente con recursos que puedan utilizarse para proteger sus intereses y en el que sea posible alcanzar un consenso. Los bienes comunes de datos pueden beneficiarse de entornos normativos más amplios diseñados específicamente para preservar los intereses del grupo (por ejemplo, mecanismos de tratados para las comunidades indígenas) frente a entidades extractivas de mayor envergadura.



# Lecturas adicionales

Chafetz, Hannah, Andrew J. Zahuranec, and Stefaan Verhulst. 2025a. “10 Data Commons for Cultural Knowledge and Preservation.” *10 Data Commons for Cultural Knowledge and Preservation*, November 24. <https://opendatapolicylab.org//articles/blog-post-10-data-commons-for-cultural-knowledge-and-preservation/>.

Chafetz, Hannah, Andrew J. Zahuranec, and Stefaan Verhulst. 2025b. *A Blueprint to Unlock New Data Commons for AI*. The Open Data Policy Lab. <https://incubator.opendatapolicylab.org/files/data-commons-for-ai-blueprint.pdf>.

Chafetz, Hannah, Andrew J. Zahuranec, and Stefaan Verhulst. 2025c. “Appendix B: Taxonomy of Data Commons Use Cases for AI.” Brooklyn, New York, March. <https://incubator.opendatapolicylab.org/files/appendix%20b.pdf>.

NYU Tandon. 2025. “The New Commons Challenge Proves the Power of Data Collectives | NYU Tandon School of Engineering.” Brooklyn, New York, October 20. <https://engineering.nyu.edu/news/new-commons-challenge-proves-power-data-collectives>.

Verhulst, Stefaan, Hannah Chafetz, and Andrew Zahuranec. 2025. “Emerging Funding and Business Models for Data Commons: A Comparison.” *Open Data Policy Lab*, October 29. <https://opendatapolicylab.org//articles/blog-post-emerging-funding-and-business-models-for-data-commons-a-comparison/>.

Verhulst, Stefaan, Hannah Chafetz, and Andrew J. Zahuranec. 2024. “Data Commons: Under Threat by or The Solution for a Generative AI Era? Rethinking Data Access and Re-Use.” SSRN Scholarly Paper No. 4836354. Social Science Research Network, May 21. <https://doi.org/10.2139/ssrn.4836354>.

Verhulst, Stefaan, Andrew J. Zahuranec, and Adam Zable. 2025. “Participatory Approaches to Responsible Data Reuse and Establishing a Social License.” In *Global Public Goods Communication: Mapping Actors, Policies, and Narratives*, edited by Sónia Pedro Sebastião and Anne-Marie Cotton. Springer Nature Switzerland. [https://doi.org/10.1007/978-3-031-90667-1\\_10](https://doi.org/10.1007/978-3-031-90667-1_10).

Verhulst, Stefaan, Andrew J. Zahuranec, Adam Zable, and Peter Addo. 2025. “Reimagining Data Governance for AI: Operationalizing a Social License for Data Reuse.” SSRN Scholarly Paper No. 5255677. Social Science Research Network, April 28. <https://doi.org/10.2139/ssrn.5255677>.

Verhulst, Stefaan, Burton Davis, and Andrew Schroeder. 2025. “Data Commons: The Missing Infrastructure for Public Interest Artificial Intelligence.” *LinkedIn*, April 29. <https://www.linkedin.com/pulse/data-commons-missing-infrastructure-public-interest-verhulst-phd-k8eec/>.

Zahuranec, Andrew J., Hannah Chafetz, and Verhulst Stefaan. 2025. “Why Responsible Data Access Will Determine the Future of AI: The Increased Importance of Data Commons.” *Open Data Policy Lab*, February 8. <https://opendatapolicylab.org//articles/blog-post-why-responsible-data-access-will-determine-the-future-of-ai-the-increased-importance-of-data-commons/>.

# Bienes comunes de datos

*Preguntas frecuentes*