



Implementación de los modelos de intercambio de datos

Estado del arte y recomendaciones de política y
gobernanza



Hechos

QUE

CONECTAN





Tabla de contenido

1. Introducción	5
2. ¿Qué se entiende por intercambio de datos?.....	6
3. Gobernanza de la infraestructura de datos	8
3.1 Niveles de gobernanza	8
3.2 Modelo de Gobernanza de la Infraestructura de Datos Colombiano (MGID)	9
3.2.1 Sobre el MGID	9
3.2.2 Roles y responsabilidades	9
3.2.3 Pilares.....	11
4. Gobernanza en modelos de intercambio de datos	13
4.1 Data Trust.....	13
4.1.1. Definiciones	13
4.1.2. Capas de la gobernanza	16
4.1.3 Tecnología	17
4.1.4 Flujos de datos.....	22
4.1.5 Oportunidades	22
4.1.6 Ejemplo internacional.....	22
4.2 Data Commons.....	23
4.2.1. Definiciones	24
4.2.2. Capas de la gobernanza	26
4.2.3 Tecnología	28
4.2.4 Flujos de datos.....	28
4.2.5 Oportunidades	29
4.2.6 Ejemplos colombianos	29
4.2.7 Ejemplo internacional.....	30
4.3 Data Marketplace	31
4.3.1. Definiciones	33
4.3.2. Data Marketplace para el sector público	34
4.3.2. Capas de la gobernanza	39
4.3.3 Tecnología	40
4.3.4 Flujos de datos.....	41
4.3.5 Oportunidades	41



4.3.6 Ejemplo internacional.....	43
4.4 Entidades líderes.....	43
5. Recomendaciones normativas e institucionales.....	44
5.1 Recomendaciones de política.....	44
5.2 Recomendaciones para los reguladores.....	45
6. Conclusiones.....	45
7. Bibliografía.....	48

Tabla de gráficos

Figura 1: Roles comunes en ecosistemas de datos.....	7
Figura 2: Jerarquía de modelos de gobernanza.....	8
Figura 3: Roles en la gobernanza de datos en Colombia.....	10
Figura 4: Los pilares para el uso efectivo de los datos.....	11
Figura 5: Las seis dimensiones de la confianza digital.....	12
Figura 6: Definiciones de Data Trusts.....	14
Figura 7: Modelo Conceptual de la plataforma de interoperabilidad X-Road.....	18
Figura 8: Modelo de consumo y explotación de Datos.....	19
Figura 9: Ejemplo de PET emergente - MPC.....	20
Figura 10: Ejemplo de analítica federada.....	21
Figura 11: Definiciones de Data Commons.....	24
Figura 12: Primer mecanismo de intercambio de datos en un mercado de datos.....	32
Figura 13: Segundo mecanismo de intercambio de datos en un mercado de datos.....	32
Figura 14: Flujo de datos en un mercado de datos.....	33
Figura 15: Definiciones de Data Marketplace.....	33
Figura 16: Características de un mercado de datos.....	35
Figura 17: Conoce el camino de los mercados emergentes de datos.....	37
Figura 18: Atributos de valor de los datos.....	38
Figura 19: Modelo de Gobernanza para un Data Marketplace.....	40



Ministerio de Tecnologías de la Información y las Comunicaciones

Carmen Ligia Valderrama Rojas - Ministra de Tecnologías de la Información y las Comunicaciones
Iván Mauricio Durán Pabón - Viceministro de Transformación Digital
Ingrid Tatiana Montealegre Arboleda – Directora de Gobierno Digital
Luisa Fernanda Medina Martínez - Subdirectora de Fortalecimiento de Capacidades Públicas Digitales
Carlos Julio León Caicedo - Líder del Equipo Datos Abiertos y Software Libre
Manuela Serrano Romero – Líder de Uso de Datos Abiertos

Departamento Nacional de Planeación

Alejandra Botero Barco - Directora General del Departamento Nacional de Planeación
Daniel Gómez Gaviria - Subdirector General Sectorial
Viviana Rocío Vanegas Barrero - Directora de Desarrollo Digital
Diana Ramírez Roa – Asesora - Dirección de Desarrollo Digital
Francisco Perea De Zubiría – Asesor - Dirección de Desarrollo Digital

Banco de Desarrollo de America Latina (CAF)

María Isabel Mejía Jaramillo - Ejecutiva Senior / Dirección de Transformación Digital

Oxford Insights

Peter Wells - Asesor experto
Pablo Fuentes Nettel - Director de proyecto
Alejandra Finotto - Investigadora principal
Kate lida - Investigadora

Versión	Observaciones
Versión 1 Mayo 2022	Implementación de los modelos de intercambio de datos: Estado del arte y recomendaciones de política y gobernanza

Comentarios, sugerencias o correcciones pueden ser enviadas al correo electrónico:
gobiernodigital@mintic.gov.co.



Este documento de la Dirección de Gobierno Digital se encuentra bajo una Licencia Creative Commons Atribución 4.0 Internacional.



1. Introducción

Desde que las sociedades actuales empezaron a hacer un uso cada vez más frecuente de herramientas tecnológicas, también incrementó el manejo de información de quienes accedían o hacían uso de estos servicios digitales. Por ello, el concepto de datos se ha transformado para determinar los parámetros de uso y gobernanza de los datos.

De acuerdo con el análisis hecho por Future Cities Canada (Future Cities Canada, 2019): “Si bien las plataformas en línea han facilitado muchos aspectos de nuestras vidas, los modelos de gobernanza de datos que emplean tienden a concentrar el acceso a los datos en manos de unas pocas grandes empresas de tecnología, lo que excluye a los ciudadanos de compartir su valor (...). Existe una necesidad apremiante de explorar nuevos modelos de gobernanza de datos que brinden a las personas cierto control sobre sus datos personales, mientras que la industria y los gobiernos trabajan para definir, proteger y desarrollar conceptos de Derechos Digitales.”

En este contexto, si bien, se debe brindar la posibilidad a los ciudadanos de fijar las reglas claras y obtener un beneficio por su información, lo ideal no sería que todo el gobierno de datos descansa sobre los hombros del sector privado, debido, precisamente, al riesgo que se correría en temas de privacidad, propiedad y uso indiscriminado de los mismos. En este orden de ideas, los nuevos usos de datos deben generar nuevas formas de gestión de datos que sean ejercidas directamente por los aparatos estatales. Sin embargo, los gobiernos utilizan métodos tradicionales de regulación y esta forma no resulta efectiva pues el ritmo del cambio tecnológico sobrepasa con gran velocidad los intentos por normativizar el uso de los datos: “Los gobiernos a menudo han tomado una actitud reactiva enfoque de la regulación y la gobernanza; primero dejan que el mercado se desarrolle libremente y luego instituyen cambios incrementales ante las amenazas y problemas. Si bien esto puede funcionar en temas que se mueven lentamente, tales enfoques son insuficientes para hacer frente al desafío de la gobernanza de datos e IA, donde los nuevos modelos de negocio pueden escalar muy rápidamente, llegando a millones de consumidores en todo el mundo en poco tiempo” (Future Cities Canada, 2019). En ese sentido, Data Trusts, Data Commons y Data Marketplaces son nuevas formas de gestionar los datos.

A partir de lo mencionado anteriormente y en el marco del CONPES 3975 de 2019 “Política Nacional para la Transformación Digital e Inteligencia Artificial” (CONPES, 2019), ¹ se realizó la primera versión de los “Lineamientos para Data Trust y Data

¹ El CONPES 3975 establece que el MinTIC tiene como compromiso para 2020 Definir los lineamientos para la implementación de modelos de Data Trusts, Data Commons, y/o Data Marketplaces en Colombia y la regulación necesaria para su operación.



Commons”. No obstante, a partir de la publicación del CONPES 4023 “Política para la reactivación, la repotenciación y el crecimiento sostenible e incluyente: Nuevo compromiso por el futuro de Colombia”, se identificó la necesidad de actualizar el documento.

Dicho lo anterior, se han creado dos documentos: 1) “**Implementación de los modelos de intercambio de datos: Estado del arte y recomendaciones de política y gobernanza**”; 2) Pasos para pilotear proyectos de intercambios de datos Data Trust, Data Commons y Data Marketplace. Dichos reportes se diferencian de la versión anterior al incorporar información sobre los mercados de datos —modelo de gobernanza que recientemente ha ganado relevancia a nivel internacional (The Data Economy Lab, 2022). El presente documento corresponde al primero de los mencionados y brinda información conceptual reciente y ejemplos actualizados sobre casos de uso (en Colombia y otros países). Asimismo identifica oportunidades generadas por el intercambio de datos en general y por cada uno de los modelos. Finalmente, se adiciona información pertinente al Modelo de Gobernanza de Infraestructura de Datos Colombiano. Es importante destacar que se busca mantener un enfoque general que permita aplicar el contenido del documento a distintos sectores de la economía colombiana.

2. ¿Qué se entiende por intercambio de datos?

El acceso y el intercambio de datos es una de las fases del ciclo de vida de los datos, y se describe como cualquier **mecanismo técnico y legal a través del cual una entidad** (individuo u organización) **puede acceder e intercambiar de manera efectiva los datos digitales**, con el fin de favorecer la reutilización de estos últimos.²

El intercambio de datos de manera general implica la interacción entre: i) los **proveedores** de datos (creadores de datos, intermediarios o vendedores de datos); ii) los **consumidores** de datos (recopilación de los datos, puede ser a través de la compra, para distintos propósitos); iii) los **proveedores de servicios** de intercambio de datos (proveen infraestructuras tecnológicas, gestionan y tratan los datos en actividades de preparación, administración y anonimización); y iv) la **autoridad que regula o habilita** el intercambio de los datos, usualmente a través de la expedición de normativas, lineamientos técnicos y definición de estándares.³

² OCDE. (2020). Enhancing access to and sharing of data: Reconciling risks and benefits for Data Re-use across Societies. Disponible en: <https://www.oecd-ilibrary.org/sites/276aaca8-en/index.html?itemId=/content/publication/276aaca8-en>

³ C4RI Colombia. (2020). Economía de Intercambio de Datos. Disponible en: <https://c4ir.co/documentos/>



De manera general, los ecosistemas de datos se componen por las redes internas y externas que permiten acceder a los datos, utilizarlos y compartirlos. Es imprescindible tener un buen ecosistema de datos para poder planear relaciones técnicas y organizativas a la hora de prestar un servicio de intercambio de datos. Los siguientes roles son comunes entre los distintos ecosistemas de datos⁴:

Figura 1: Roles comunes en ecosistemas de datos

Guardianes de los datos	Responsable de recopilar, gestionar y garantizar el acceso a un conjunto de datos; puede incluir la provisión de infraestructura, la gobernanza de los datos, etc.
Individuos u organizaciones	Personas u organizaciones a las que se refieren los datos o que se ven afectadas por su uso.
Colaboradores	Personas u organizaciones que contribuyen o ayudan a conservar un conjunto de datos; pueden hacerlo conscientemente, utilizando herramientas y marcos proporcionados por un administrador de datos, o sin saberlo a través del uso de un servicio.
Reguladores	Crear las políticas y los marcos legislativos en los que operan los demás actores.
Intermediarios	Proporcionar servicios de valor añadido que enriquezcan un conjunto de datos.
Agregadores	Tipo de intermediario. Agrupa conjuntos de datos provenientes de muchas fuentes.
Creadores (o reutilizadores)	Utilizar los datos para crear información, en forma de productos y servicios, análisis y perspectivas, o historias y visualizaciones.
Beneficiarios	Personas u organizaciones que se benefician del ecosistema de datos al tomar mejores decisiones informadas por el uso de productos y servicios, junto con su propia experiencia y conocimiento.
Investigadores	Tipo de creador. Utiliza los datos con fines de investigación.
Responsables de política pública	Crear principios y métricas para generar y evaluar resultados.

Fuente: ODI

⁴ ODI. (2022). Data ecosystem mapping: plotting the journey from data to value for your business. Disponible en: https://theodi.org/event_series/data-ecosystem-mapping-plotting-the-journey-from-data-to-value-for-your-business/

3. Gobernanza de la infraestructura de datos

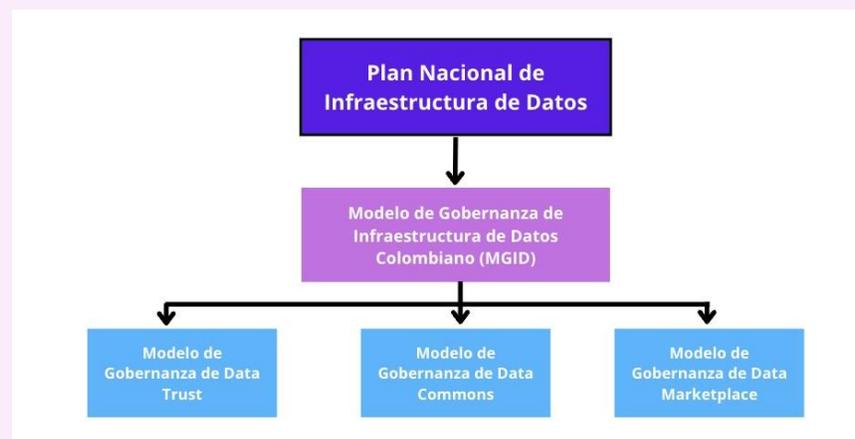
3.1 Niveles de gobernanza

Un modelo de gobernanza para el intercambio de datos entre el sector público y privado debe plantearse alrededor de dos niveles. El primero se refiere a **estrategias a nivel nacional** que minimicen las barreras de la compartición y establezcan políticas en torno a la gobernanza general de datos. El segundo relacionado con el desarrollo de **mecanismos de intercambios de datos**.

En el primer nivel de gobernanza se proponen la definición de estrategias para promover la confianza y la cultura de intercambio de datos, definición de estándares que faciliten el intercambio de datos entre el sector público y el sector privado, la promoción de espacios de intercambios de datos a nivel de plataformas tecnológicas que compartan una arquitectura común (como por ejemplo, mercados de datos) y que se constituyan en entornos seguros para el tratamiento de datos y la definición de registros base o datos maestros que habiliten el intercambio de datos entre entidades del Estado y también con el sector privado.

Es importante recalcar que además del Modelo de Gobernanza de la Infraestructura de Datos Colombiano, **cada mecanismo de intercambio de datos debe contar con un esquema de gobierno propio que se adapte a las condiciones específicas del caso de uso (segundo nivel)**⁵. Dichas estipulaciones, sin embargo, deberán estar alienadas al MGID.

Figura 2: Jerarquía de modelos de gobernanza



Fuente: Elaboración propia

⁵ Para explorar el segundo nivel de gobernanza, referirse a las secciones 4.1, 4.2 y 4.3.



3.2 Modelo de Gobernanza de la Infraestructura de Datos Colombiano (MGID)

3.2.1 Sobre el MGID

A la hora de definir la gobernanza del Data Commons, Data Trust, y Data Marketplace, es necesario tener en cuenta el marco de gobernanza general de Colombia (primer nivel). Dicho mecanismo sirve para brindar certeza sobre la toma de decisiones, responsabilidades y atribuciones de los actores involucrados en un determinado régimen económico. En ese sentido, el gobierno colombiano cuenta con el **Modelo de Gobernanza de la Infraestructura de Datos Colombiano (MGID)**.

El MGID es un “sistema de elementos políticos, técnicos, legales y organizacionales que trabajan de forma dinámica y coordinada para la generación de valor público a través de los datos” (MGID, 2022). El modelo presenta un “esquema de responsabilidad distribuida, donde intervienen diversos actores para promover la democratización de la infraestructura de datos con el fin de que los diversos actores del ecosistema de datos, incluida la ciudadanía, se vean beneficiados, y promueva el desarrollo en un contexto de creciente digitalización” (MGID, 2022).

El modelo surge a partir de la coordinación con actores del ecosistema de datos colombiano, así como a niveles regional e internacional. Este Modelo se sustenta en los principios de gobernanza previamente definidos en el **Plan Nacional de Infraestructura de Datos**, con la finalidad de garantizar la continuidad y efectividad de una misma visión de la infraestructura de datos del Estado Colombiano (MGID, 2022):

- Calidad de los datos;
- Fácil búsqueda, accesibilidad, interoperabilidad y reutilización;
- Seguridad y protección de los datos;
- Privacidad por diseño y por defecto;
- Confianza pública y gestión ética de los datos;
- Estandarización e interoperabilidad;
- Sectorización estratégica.

3.2.2 Roles y responsabilidades

Dentro del MGID propuesto para Colombia, se han definido distintos roles para personas según su conocimiento técnico y liderazgo. Este componente incluye a los siguientes roles e instancias que adoptarán funciones y responsabilidades sobre planeación estratégica, rendición de cuentas, coordinación e implementación de políticas de datos, cuya descripción y responsabilidades se incluyen a continuación:



Figura 3: Roles en la gobernanza de datos en Colombia

Líder Nacional de los Datos o coordinador nacional de datos: rol encargado de encabezar y liderar la estrategia de la Infraestructura de Datos al más alto nivel. Su principal responsabilidad es promover la alineación interinstitucional para la explotación de la infraestructura de datos en el sector público y el uso de los datos en la toma de decisiones.

Chief Data Officers: rol encargado de coordinar los procesos de cambio dentro de las entidades públicas, así como de aumentar el aprovechamiento de la infraestructura de datos y las estrategias que se definan en el marco de PNID y posterior a su implementación. Su principal responsabilidad es asegurar el cumplimiento de los lineamientos, directrices y guías que el gobierno nacional publique respecto al uso, aprovechamiento y explotación de datos.

Gerentes de datos: rol con responsabilidades directamente relacionadas con la creación, procesamiento, almacenamiento, intercambio, uso y análisis, archivo y preservación de los datos en todas las áreas y niveles del gobierno. Sus responsabilidades principales son de cumplir con los estándares, lineamientos y guías que el gobierno defina respecto al ciclo de vida de los datos, y de mantener actualizada la información respecto a los diccionarios de datos que se generen.

Comité Técnico de Datos: el comité técnico se compone de representantes del gobierno nacional, representantes de sociedad civil, representantes del sector privado o de la industria, y representantes del sector académico. Sus responsabilidades principales son fomentar la inclusión e integración de agendas de los diversos sectores y stakeholders que intervienen en la infraestructura de datos del estado Colombiano

Equipo Técnico para la Implementación: un grupo de especialistas en ciencia de datos que se encargará de capacitar, asesorar y acompañar a los gerentes de datos. Sus responsabilidades principales son de brindar capacitación y asesoría técnica para la implementación del PNID, y de las estrategias que deriven a mediano y largo plazo, y de proveer acompañamiento técnico a los involucrados en los casos de uso de datos.

Fuente: Elaboración propia, basada en MGID, 2022.

3.2.3 Pilares

Figura 4: Los pilares para el uso efectivo de los datos



Fuente: Elaboración propia basada en MGID, 2022.

Según el Modelo de Gobernanza de Infraestructura de Datos de Colombia (MGID, 2022), existen **cinco condiciones necesarias para el uso eficaz de los datos**. Estas son: interoperabilidad, tecnología ética y responsable, seguridad y protección de los datos, capacidades digitales y alta calidad de los datos, y se exploran en el siguiente gráfico. Es imprescindible aplicar estos pilares en el desarrollo de un Data Commons, Data Trust y Data Marketplace.

A estos pilares se les deben añadir las siguientes seis dimensiones de confianza digital: auditabilidad, rendición de cuentas, ética, equidad, seguridad y transparencia. Estas son clave a tener en cuenta en el desarrollo de un Data Commons, Data Trust y Data Marketplace, ya que son imprescindibles a la hora de incrementar confianza entre distintos actores y fomentar el diálogo entre estos (WEF, 2021).

Figura 5: Las seis dimensiones de la confianza digital



Fuente: PwC Colombia and the World Economic Forum, 2021

4. Gobernanza en modelos de intercambio de datos

Si bien el MGID presenta los lineamientos generales para el gobierno de datos en Colombia, cada modelo de gobernanza de datos presenta sus propias particularidades. Un **mecanismo de intercambio de datos deberá plantear su propio modelo de gobernanza que se adapte a lo estipulado por el MGID y, al mismo tiempo, se adapte a las necesidades del tipo de institución y a las condiciones específicas del sector**. A continuación información clave sobre el tema.

4.1 Data Trust

Los Data Trust son un modelo de gobierno de datos. Se diferencian de los Data Commons y los Data Marketplaces en que su modelo de gobernanza se basa en un consejo de administración que tiene la responsabilidad de proteger y compartir los datos de forma ética.

Propuestos por primera vez por Lilian Edwards en 2004⁶, los Data Trusts son un **área activa de investigación** por parte de múltiples organizaciones de investigación y política de datos.

4.1.1. Definiciones

La Open Data Institute - ODI define los Data Trust del siguiente modo (Open Data Institute, 2020): "Un fideicomiso de datos proporciona una administración fiduciaria e independiente de los datos. Los fideicomisos de datos son un enfoque para cuidar y tomar decisiones sobre los datos de una manera similar a la que se han utilizado para cuidar y tomar decisiones sobre otras formas de activos en el pasado [...] Como administrador de los datos, un fideicomiso de datos puede decidir quién tiene acceso, en qué condiciones y en beneficio de quién".

Por lo general, los Data Trust se fundamentan como una **estructura legal que permite la custodia independiente de algunos datos para el beneficio de un grupo de organizaciones o personas**, que puede ser la creación de empresas, la investigación de enfermedades o el empoderamiento de comunidades. Establecen una **clara responsabilidad "fiduciaria"** para las empresas u organizaciones que custodian los

⁶ Lilian Edwards. (2004). The Problem With Privacy. Disponible en: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1857536

datos. Es un instrumento que utiliza la infraestructura del derecho privado sin depender excesivamente de la acción del gobierno.

En varios países se está trabajando en la creación de pilotos de fideicomisos de datos y en la modificación de los marcos legales para facilitar su creación. No obstante, aún no se presentan metodologías ampliamente aceptadas para su desarrollo. **Existen pocos ejemplos activos y, en consecuencia, la evidencia sobre los resultados de los *Data Trust* es escasa.** Hasta hoy, no es muy claro si el concepto tiene potencial para aumentar el acceso a los datos y mantener la confianza de las distintas partes interesadas.

Los **fideicomisos de datos son un concepto emergente** que ha sido definido de diversas maneras por diferentes organizaciones y personas. Las definiciones se ven afectadas por los objetivos de las diferentes organizaciones (por ejemplo: maximizar el crecimiento económico, maximizar la disponibilidad de datos o abordar los desequilibrios de poder entre los diferentes actores).

A continuación un listado de definiciones de Data Trusts:

Figura 6: Definiciones de Data Trusts

Organización	Definición
Data Economy Lab ⁷	Permiten que los usuarios hagan de sus datos un bien común (o derechos de datos) para negociar colectivamente sus condiciones de uso, en un marco de fuertes obligaciones fiduciarias.
Open Data Institute ⁸	Un <i>Data Trust</i> proporciona una administración independiente y fiduciaria de los datos.
Mozilla ⁹	Un fideicomiso de datos es una relación jurídica en la que un fideicomisario administra los derechos de los datos en interés exclusivo de un beneficiario o de un grupo de beneficiarios.

⁷ The Data Economy Lab. (2022). Data Stewardship models. Disponible en: <http://tool.thedataeconomylab.com/our-data-models>

⁸ ODI. (2020). Data trusts in 2020. Disponible en: <https://theodi.org/article/data-trusts-in-2020/>

⁹ Mozilla. (2020). What Does it Mean? | Shifting Power Through Data Governance. Disponible en: <https://foundation.mozilla.org/en/data-futures-lab/data-for-empowerment/shifting-power-through-data-governance/#what-is-a-data-trust>



Global Partnership on AI ¹⁰	Un tipo de institución de datos que ayuda a individuos o grupos a poner en común recursos, encargando a un "fiduciario" independiente la gestión de esos recursos en beneficio de los miembros del fideicomiso. Caracterizados por: permitir la innovación impulsada por los datos en beneficio social y económico, creando un entorno de confianza para el intercambio de datos; equilibrar las asimetrías de poder en los intercambios de datos, alentando y capacitando a los creadores de los datos para que desempeñen un papel activo en el establecimiento de las condiciones de uso de los datos -y la distribución del valor que crean- y proporcionando una plataforma para la negociación colectiva; y anticipar, prevenir y gestionar las vulnerabilidades asociadas al uso de los datos, mediante una administración profesional de los mismos.
Sean Mcdonald (Digital Public) ¹¹	Los fideicomisos son instrumentos jurídicos que designan a un administrador (fiduciario) para que gestione un activo con un fin -como la conservación de la tierra o la maximización del valor- en nombre de un beneficiario o beneficiarios que poseen el activo. Los fideicomisos de datos son fideicomisos legales que gestionan los datos, o los derechos a los datos.
Data Trusts Initiative ¹²	Un fideicomiso de datos es un mecanismo para que las personas tomen los derechos de datos que establece la ley y los agrupen en una organización -un fideicomiso- en la que los fideicomisarios toman decisiones sobre el uso de los datos en su nombre.
UK Department for Work and Pensions ¹³	Un marco de confianza para la colaboración digital.
Brighthive ¹⁴	Un marco legal, técnico y de gobernanza que permite a un colectivo de organizaciones conectar de forma segura sus respectivas fuentes de datos y crear nuevos recursos de datos compartidos que beneficien a cada miembro

Fuente: Elaboración propia.

¹⁰ OECD.AI. (2021). Enabling data sharing for social benefit through data trusts at the Global Partnership on AI, GPAI. Disponible en: <https://oecd.ai/en/work/data-sharing-data-trusts>

¹¹ CIGI. (2019). Reclaiming Data Trusts. Disponible en: <https://www.cigionline.org/articles/reclaiming-data-trusts/>

¹² Data Trusts UK. (2022). The Data Trusts Initiative. Disponible en: <https://datatrusts.uk/about>

¹³ DWP Digital. (2022). DWP Digital is improving cross-government data sharing. Disponible en: <https://dwpdigital.blog.gov.uk/2022/02/17/dwp-digital-is-improving-cross-government-data-sharing/>

¹⁴ Brighthive. (2022). The Use of Data Trusts to Improve Impact (Responsibly). Disponible en: <https://brighthouse.io/the-use-of-data-trusts-to-improve-impact-responsibly/>



4.1.2. Capas de la gobernanza

El modelo de gobernanza de los Data Trust se basa en un consejo de administración que tiene la responsabilidad de proteger y compartir los datos de forma ética. Es una estructura legal que permite la custodia independiente de algunos datos para el beneficio de un grupo de organizaciones o personas. Establecen una clara responsabilidad “fiduciaria” para los que custodian los datos. Un Data Trust es un instrumento que utiliza la infraestructura del derecho privado sin depender excesivamente de la acción del gobierno.

Un Data Trust debería contener al menos los siguientes componentes:

- Un acuerdo claro y bien definido entre fideicomitente y fiduciario.
- Activos de datos como identificadores, registros y conjuntos de datos.
- Los estándares y tecnologías utilizados para seleccionar y proporcionar acceso a esos activos de datos.
- La orientación y las políticas que informan el uso, la seguridad y la gestión de los activos de datos y la propia infraestructura de datos.
- Las organizaciones que gobiernan la infraestructura de datos.
- Las comunidades involucradas en la contribución o el mantenimiento de la infraestructura de datos, y aquellos que se ven afectados por las decisiones que se toman al utilizarla.
- Estricta auditoría y traza del procesamiento de datos.
- Acceso de información ágil a los científicos de datos con transparencia y responsabilidad.

En varios países se está trabajando en la creación de pilotos de fideicomisos de datos y en la modificación de los marcos legales para facilitar su creación. No obstante, aún no se presentan metodologías ampliamente aceptadas para su desarrollo. Existen pocos ejemplos activos y, en consecuencia, la evidencia sobre los resultados de los Data Trust es escasa. Hasta hoy, no es muy claro si el concepto tiene potencial para aumentar el acceso a los datos y mantener la confianza de las distintas partes interesadas. Al mismo tiempo, cualquier impacto se verá afectado por el propósito de la organización que lo ejecuta, cómo se lleva a cabo y el ecosistema más amplio dentro del cual opera.

Respecto a los beneficios, autores como Suzor y Gray (Suzor & Gray, 2018) comentan: “Un fideicomiso de datos es un modelo de conocimiento común para el intercambio de datos que puede empoderar y crear valor para sus miembros. [...] El fideicomiso de datos crea un nuevo recurso para los participantes del mercado: es un recurso económico que se comparte para beneficio mutuo”.



4.1.3 Tecnología

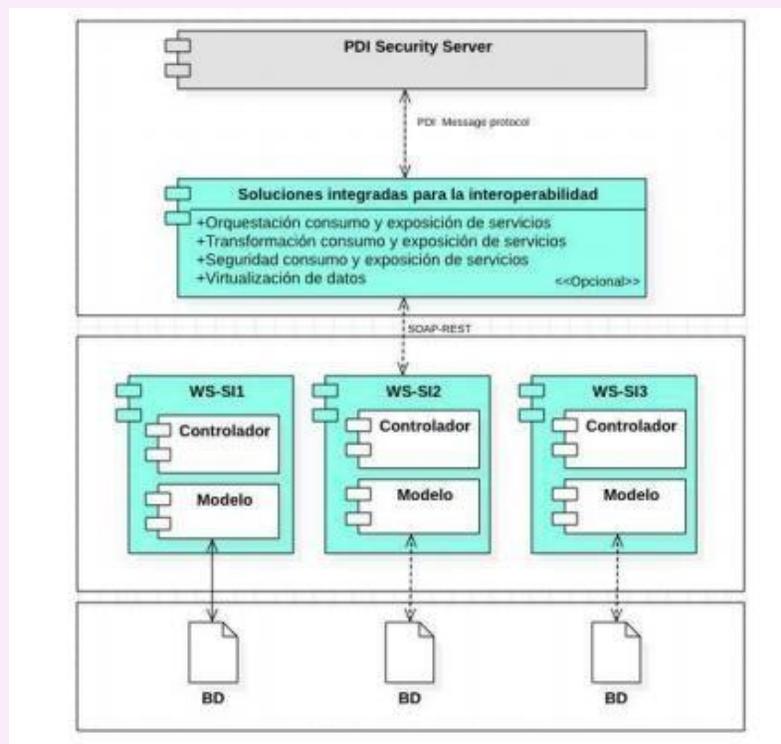
La tecnología utilizada por los Data Trusts y los Data Commons puede ser muy similar. Sin embargo, hay algunas especificaciones claves que son importantes para los Data Trusts en particular. Los Data Trusts son muy útiles para ayudar a crear confianza para intercambiar datos en un entorno en el que hay menos confianza entre los actores. Esto puede tener que ver, por ejemplo, con que los datos sean muy sensibles, como los datos personales relacionados con la salud. Debido a la naturaleza sensible de los datos, los Data Trusts pueden necesitar cumplir más requisitos de auditabilidad, verificabilidad y transparencia. Normalmente, el flujo de datos en un Data Trust se supervisa más que en un Data Commons. Se puede supervisar con sistemas de gestión de acceso, sistemas de autenticación y registros accesibles y verificables que rastreen por dónde fluyen los datos.

En los Data Trusts en los que se comparten datos sensibles y personales, existen leyes individuales de protección de datos que pueden afectar también a la tecnología utilizada en el Data Trust. Existen distintas plataformas de interoperabilidad que pueden ser útiles para facilitar la confianza en el intercambio de datos dentro de un Data Trust. Un ejemplo es la plataforma X-Road, la cual permite a los individuos ver sus flujos de datos personales (MinTIC, 2019).

X-Road funciona de esta manera:

Cada administrador o recolector de datos debe disponer de un servidor seguro para comunicarse a través del protocolo X-Road, como se puede ver en la imagen.

Figura 7: Modelo Conceptual de la plataforma de interoperabilidad X-Road



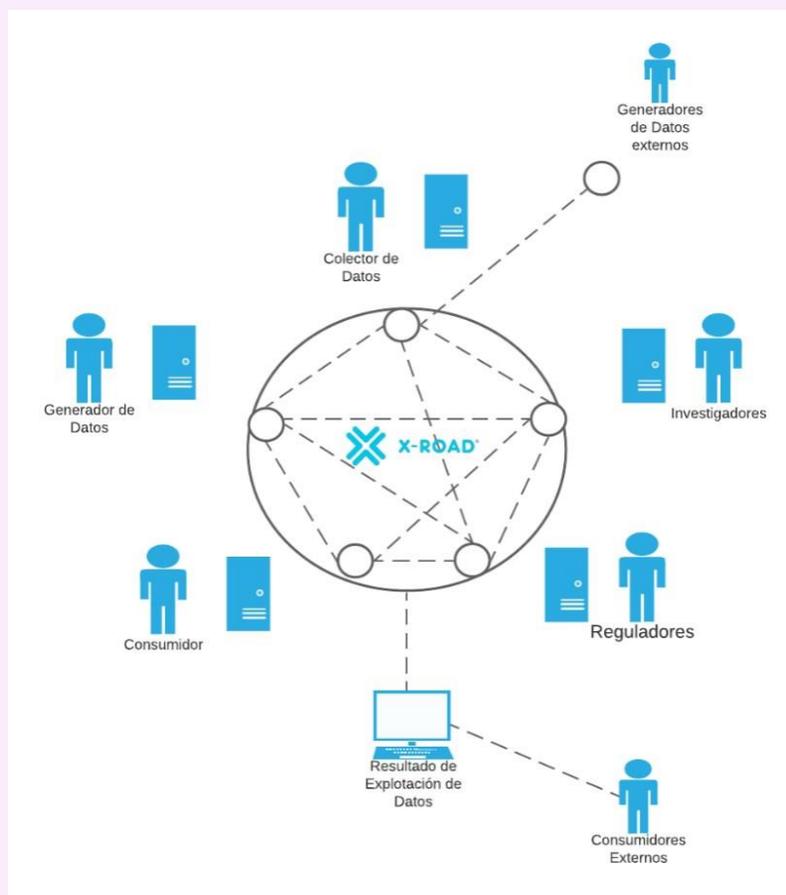
Fuente: MinTIC, 2019

De este modo, toda la arquitectura convergerá en un modelo descentralizado. En este modelo todos los servicios se construyen entre uno o varios recolectores o administradores de datos y se exponen para el intercambio seguro mediante X-Road para su consumo por quienes requieren y retroalimentan la información, quienes a su vez, crearían servicios con la información producto de su análisis, así, se construirían diferentes redes con los modelos de datos diseñadas e implementadas con propósitos específicos, por ejemplo: una investigación sobre algún tema referente a la salud, transporte o agricultura, entre otros.

En el modelo presentado a continuación, se tiene en cuenta que existen generadores de datos que no podrían adquirir un servidor de seguridad X-Road (como, por ejemplo, personas naturales), por lo tanto, sus datos deben ser recopilados por los colectores, además, la red descentralizada debe generar un resultado de explotación el cual podrá ser consumido por usuarios externos.

Un Data Trust puede contar con esta estructura pero sus resultados pueden o no publicarse (según sus propias políticas). Los Data Trust deberán compartir la información bajo el mismo protocolo de intercambio de mensajes a través de la herramienta X-Road y así se convertiría en otro colector de datos.

Figura 8: Modelo de consumo y explotación de Datos



Fuente: Elaboración propia

En este contexto, los actores dentro de la red son los siguientes:

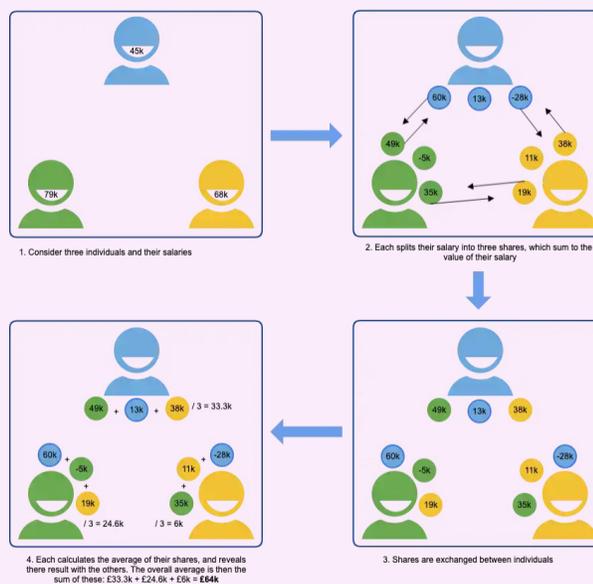
- **Generador de datos:** es quien genera el dato en primer lugar, por ejemplo, personas naturales que generan sus datos de contacto.
- **Colector de datos:** es quien agrupa la información de los generadores en bases de datos, por ejemplo, las entidades que comparten los conjuntos de datos.
- **Investigador:** es quien toma los datos primarios de los colectores o generadores y los agrupa, estructura, centraliza y explota sobre una investigación o propósito en particular, finalmente, generará un nuevo o nuevos conjuntos de datos con el resultado de la investigación. En este contexto, las investigaciones pueden generarse en cualquier ámbito, como: salud, agricultura, economía, movilidad, entre otras, y dichas investigaciones se podrían generar desde la academia, los ministerios u otras entidades.

- **Consumidor:** es quien obtiene los conjuntos de datos resultantes como insumo para sus procesos
- **Regulador:** es quien audita y regula el modelo de datos y proceso que generó los conjuntos de datos dentro de la red.

Los Data Trusts suelen necesitar sistemas para proteger la privacidad y la confidencialidad de la información sensible. Las PET (privacy enhancing technology) son una tecnología muy útil para ello. Algunas PETs emergentes son, por ejemplo, los protocolos MPC (multi-party computation). Un "protocolo de cálculo multipartito" (MPC) permite calcular conjuntamente una función que recibe información de varias partes, mientras que cada parte participante mantiene su información en secreto para los demás. Para ello, se fragmentan los datos en varios nodos de la red, de modo que cada nodo alberga un "fragmento ininteligible" de datos; la inspección de un fragmento individual no revela información sobre los datos originales. Cada nodo calcula una función en su fragmento y los resultados se agregan a un resultado final (Centre for Data Ethics and Innovation, 2021).

Para comprender de forma intuitiva el MPC, pensemos en un grupo de empleados que quieren determinar su salario medio, sin revelar sus salarios individuales a los demás. La siguiente figura ilustra cómo se puede implementar el MPC utilizando la compartición de secretos aditivos, basándose en las propiedades matemáticas básicas de la adición para dividir el cálculo entre las partes de forma que los salarios sigan siendo confidenciales. El resultado correcto se obtiene combinando los resultados de los cálculos de cada parte (Centre for Data Ethics and Innovation, 2021).

Figura 9: Ejemplo de PET emergente - MPC

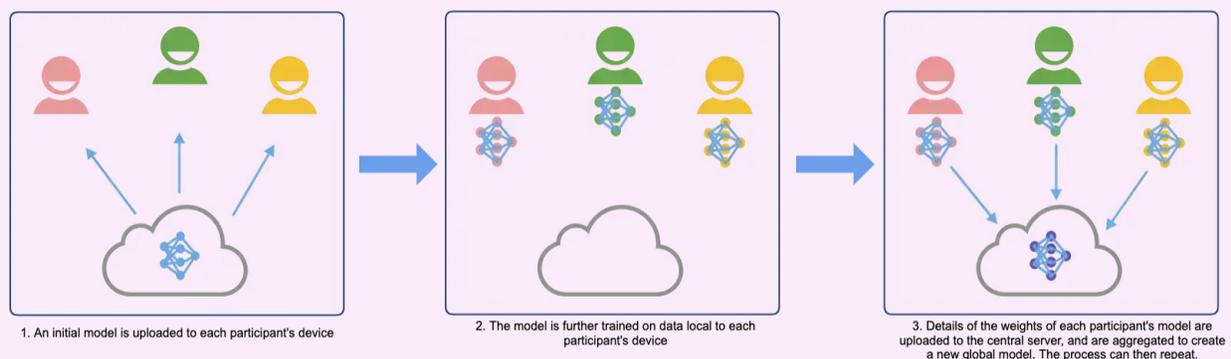


Fuente: Centre for Data Ethics and Innovation, 2021

Otro ejemplo interesante de PET emergente es la Analítica Federada. La analítica federada es un paradigma para ejecutar un programa informático con datos descentralizados. Se trata de que una entidad o persona cargue el programa en el servidor o dispositivo en el que se encuentran los datos, lo ejecute en el servidor o dispositivo con los datos in situ y comunique los resultados a la parte que lo originó. De este modo, no se revela ningún dato directamente a la entidad o persona (Centre for Data Ethics and Innovation, 2021).

Un subconjunto de la analítica federada es el aprendizaje federado, que consiste en entrenar un modelo de aprendizaje automático en conjuntos de datos distribuidos. La idea es entrenar modelos locales directamente en los dispositivos de los usuarios utilizando datos locales, y luego que los dispositivos compartan los pesos del modelo resultante entre sí, para que se determine un nuevo modelo global. Puede tratarse de un aprendizaje federado centralizado, en el que un servidor central se encarga de coordinar las acciones de los dispositivos participantes (como se muestra a continuación), o de un aprendizaje federado descentralizado, en el que los dispositivos participantes se coordinan entre sí. En cualquiera de los dos casos, la característica principal es que los datos del usuario nunca salen del dispositivo, ya que sólo se comunican los pesos del modelo (Centre for Data Ethics and Innovation, 2021). Un ciberataque podría inferir la información del usuario a partir de estos pesos, y a menudo se incorporan mecanismos adicionales para mitigar esto.

Figura 10: Ejemplo de analítica federada



Fuente: Centre for Data Ethics and Innovation, 2021

Otro aspecto a tener en cuenta respecto de los Data Trust descentralizados construidos con tecnología blockchain (Mills, 2019), es que permite que se valide el origen de los datos. De esta forma, los generadores de datos agregarían sus datos usando una cuenta encriptada, protegiendo su anonimato, mientras que todas las cargas y descargas de datos se registran en un blockchain de acceso público, estableciendo responsabilidades. Los datos solo se volverían a identificar después del permiso del generador de datos y podrían moderarse mediante contratos inteligentes para limitar el tiempo de acceso.



Finalmente, el sistema descentralizado, asegura que ninguna entidad individual controle una gran cantidad de datos, lo que reduce el riesgo de filtraciones de datos y formación de monopolios.

4.1.4 Flujos de datos

En los Data Trusts, Data Commons y Data Marketplaces, los datos pueden estar centralizados o descentralizados. El factor más importante para decidir sobre el flujo de datos es tiene que ver con el caso de uso específico (no el modelo de gobernanza elegido). En los Data Trusts, es más útil pensar en el flujo de los derechos de los datos que en el flujo de los propios datos.

El flujo de derechos de datos en un Data Trust sigue este modelo. Hay fideicomisarios que reciben el derecho a tomar decisiones sobre los conjuntos de datos. Los fideicomisarios y el consejo de administración tienen derecho a decidir cómo podrán utilizar estos conjuntos de datos.

4.1.5 Oportunidades

El impacto de un Data Trust podría incluir lo siguiente:

- Cambios futuros en la legislación de datos para apoyar la **gobernanza colectiva de datos**. Si los enfoques de gobernanza colectiva de datos (como los fideicomisos de datos) resultan ser un mecanismo útil, ser uno de los primeros gobiernos en adoptar estos enfoques en la legislación podría tener importantes beneficios a largo plazo.
- **Mayor acceso a los datos**. Algunos modelos de fideicomiso de datos tendrán el propósito explícito o la consecuencia de aumentar el acceso a los datos. Este mayor acceso podría conducir a otros beneficios, como una mayor responsabilidad, una mayor innovación o mejoras en la toma de decisiones y la eficiencia.
- Organizaciones y productos digitales más confiables. Un argumento en favor de los fideicomisos es que los modelos de gobernanza colectiva tienen potencial para **fortalecer vínculos entre usuarios y proveedores**. Esto puede dar lugar a mayores niveles de colaboración y productos digitales más fiables.

4.1.6 Ejemplo internacional



PLACE

PLACE es una organización sin ánimo de lucro que quiere democratizar el acceso a los datos cartográficos. La mayoría de los datos cartográficos están concentrados en pocas empresas grandes y no son igualmente accesibles. PLACE quiere hacer más accesibles los datos cartográficos para que puedan utilizarse en beneficio de las sociedades. PLACE crea y ofrece imágenes hiperlocales precisas de la superficie terrestre. Para proteger los datos que recopilan y utilizan, PLACE ha creado un fideicomiso con sede en el Reino Unido, que guarda todos los datos y licencias gubernamentales que PLACE ha recibido. PLACE cuenta con un consejo de administración que tiene la responsabilidad fiduciaria de la organización. Otro consejo supervisa el fideicomiso. Para utilizar los datos de PLACE, las organizaciones deben convertirse en miembros del fideicomiso. Los miembros firman un contrato que incluye obligaciones relativas al uso ético de los datos. Las organizaciones que deseen utilizar los datos de PLACE deben pagar una cuota de afiliación.

4.2 Data Commons

El concepto de *Data Commons* se basa en la idea de que los **datos deben ser útiles para una comunidad o para el público en general**. Parte del argumento de que se debe garantizar que todos los agentes cuenten con las mismas condiciones en cuanto al acceso a los datos. Actualmente, existen dos principales categorías de Data Commons: **investigación** y **recurso común (*common pool*)**.

La categoría de investigación, se ha puesto en marcha mediante la implementación de mecanismos de compartición de datos en áreas como la salud pública y la labor académica en ciencias sociales. Tiene la finalidad de eliminar barreras para el intercambio de información, lo que, eventualmente, podría llevar a la generación de conocimiento.

Por otro lado, la categoría de *data pool*, clasificar los datos como recursos comunes y busca imitar modelos de gobernanza utilizados en actividades económicas como la pesca o la silvicultura (donde las materias primas no están sujetas a estipulaciones de propiedad privada). Esta última idea se basa en que los individuos de una comunidad deben tener el mismo nivel de acceso de los datos, sin ningún tipo de candado.



4.2.1. Definiciones

Las definiciones de los Data Commons se dividen en dos grupos principales. El primer grupo, donde se encuentran las definiciones de *Mozilla Data Futures Lab*, “En un Data Commons los datos se comparten como un recurso común (Mozilla Foundation, 2022). A menudo, van acompañados de un alto grado de propiedad y liderazgo de la comunidad y tienen una causa de bien público” y *Data Commons Barcelona*, “Un Data Commons sólido va más allá de los datos abiertos. Atiende no sólo al potencial del acceso abierto a los datos sino también a las condiciones de su producción; al poder sobre ellos y a la protección de los mismos; a su gobernanza; así como a la responsabilidad social por sus impactos”, se inspiran en el trabajo de la economista Elinor Ostrom (1990) —ganadora del premio Nobel, por su investigación sobre la **gobernanza de recursos comunes**, como la pesca y los bosques, y su aplicación a los datos.

Otras definiciones, como la de Robert Grossman, “Infraestructura interoperable que coordina los datos, el almacenamiento y la informática con herramientas de análisis comunes” y la Declaración de Berlín sobre la infraestructura nacional de datos de investigación, “Un *Data Commons* de investigación es una infraestructura virtual global ampliable para aprovechar la participación de los usuarios y la investigación colaborativa basada en datos. Esto incluye, por ejemplo, servicios conjuntos en la nube, acceso a la potencia de cálculo y espacios de trabajo colaborativos, y una infraestructura común de autenticación y autorización”, se centran en los **elementos técnicos necesarios** para que múltiples personas y organizaciones compartan datos y colaboren en el análisis de éstos. Estas definiciones suponen implícitamente la existencia de procedimientos estándar de gobernanza de la investigación, como códigos profesionales y comités de ética.

A continuación un listado de definiciones de Data Trusts:

Figura 11: Definiciones de Data Commons

Organización	Definición
Mozilla Data Futures Lab ¹⁵	En un Data Commons los datos se comparten como un recurso común. A menudo, van acompañados de un alto grado de propiedad y liderazgo de la comunidad y tienen una causa de bien público

¹⁵ Mozilla. (2020). What Does it Mean? | Shifting Power Through Data Governance. Disponible en: <https://foundation.mozilla.org/en/data-futures-lab/data-for-empowerment/shifting-power-through-data-governance/#what-is-a-data-commons>



Data Commons Barcelona ¹⁶	Un Data Commons sólido va más allá de los datos abiertos. Atiende no sólo al potencial del acceso abierto a los datos sino también a las condiciones de su producción; al poder sobre ellos y a la protección de los mismos; a su gobernanza; así como a la responsabilidad social por sus impactos.
Aotearoa Zealand Commons ¹⁷	Es la forma en que una comunidad comparte sus datos. El principio de los bienes comunes implica un recurso que pertenece a toda la comunidad y que se comparte por igual entre todos los miembros de la misma, con sujeción a las normas de la comunidad.
Common Wealth ¹⁸	Las plataformas y los datos deben conceptualizarse como servicios públicos y activos con nuevas formas públicas y comunes de propiedad y gobernanza.
Nesta ¹⁹	Una nueva forma de bien público que permite a las personas compartir sus datos, de forma anónima, con fines específicos o de beneficio social.
Data Economy Lab ²⁰	La idea de los <i>Data Commons</i> se basa en la premisa de que el valor derivado de los datos pertenecientes a un colectivo debe ser controlado por el mismo colectivo y beneficiarlo principalmente.
Robert Grossman, Allison Heath, Mark Murphy ²¹	Infraestructura interoperable que coordina los datos, el almacenamiento y la informática con herramientas de análisis comunes.
Berlin Declaration on NFDI (national research data infrastructure) Cross-	Un <i>Data Commons</i> de investigación es una infraestructura virtual global ampliable para aprovechar la participación de los usuarios y la investigación colaborativa basada en datos. Esto incluye, por ejemplo, servicios conjuntos en la nube, acceso a la potencia de cálculo y espacios de trabajo colaborativos, y una infraestructura común de

¹⁶ Data Commons Barcelona (2019). Our Vision. Disponible en: <https://datacommons.barcelona/our-vision/>

¹⁷ The Data Commons Blueprint (2020). Disponible en: <https://web.archive.org/web/20210119013750/https://datacommons.org.nz/>

¹⁸ Common Wealth. (2020). A Common Platform: Reimagining Data and Platforms. Disponible en: <https://www.common-wealth.co.uk/reports/common-platform-tech-utility-antitrust#chapter-7>

¹⁹ Nesta. (2022). Unlocking the value in data as a commons. Disponible en: <https://www.nesta.org.uk/feature/four-future-scenarios-personal-data-economy-2035/unlocking-the-value-in-data-as-a-commons/>

²⁰ Aapti Institute. (2020). Data sharing for good in business and governance: Principles behind the process. Diponible en: <https://aapti.medium.com/data-sharing-for-good-in-business-and-governance-principles-behind-the-process-7f7b853c99ab>

²¹ Robert Grossman et al. (2016). A Case for Data Commons: Towards Data Science as a Service. Disponible en: <https://arxiv.org/abs/1604.02608>



Fuente: Elaboración propia

4.2.2. Capas de la gobernanza

Su modelo de gobernanza se basa en la idea de que debe haber igualdad de acceso a los datos. Los miembros de un Data Commons, por tanto, comparten los datos entre ellos de forma igualitaria.

Funcional y técnicamente, un Data Common es similar a un Data Trust, sin embargo, presenta una diferencia sustancial en el objetivo, el cual representa un bien común y por lo tanto es un modelo generalmente público, como explica Grossman (Grossman, How Data Commons Can Support Open Science, 2019): “Los proyectos y las comunidades utilizan los datos comunes para crear recursos abiertos a fin de acelerar la tasa de descubrimiento y aumentar el impacto de los datos que alojan.”, a lo cual agrega, que los datos comunes deben respetar los principios FAIR (GO FAIR, 2016):

- **Localizable:** asignando identificadores únicos y persistentes, describiendo los datos con metadatos enriquecidos, incluyendo el identificador asignado e indexándolos en un recurso de búsqueda.
- **Accesible:** utilizando protocolos estandarizados de comunicación que sean abiertos y gratuitos. Cuando los datos no puedan ser abiertos, el protocolo debe permitir procedimientos para la autenticación y la autorización.
- **Interoperable:** los metadatos deben utilizar formatos, lenguajes y vocabularios acordados por la comunidad y contener enlaces a información relacionada mediante identificadores.
- **Reusable:** asignando metadatos con atributos que proporcionen información contextual y metadatos de información sobre su procedencia. Deben utilizar una licencia abierta y legible por ordenador y estándares que use la comunidad del dominio concreto, para permitir su reutilización.

A continuación, algunas características de los Datos Commons:

- Los datos comunes hacen que los datos estén disponibles para que estén abiertos y se pueda acceder a ellos y analizarlos con facilidad.
- A diferencia de un lago de datos o Data Lake, los datos se limpian y se brinda calidad a los datos utilizando uno o más modelos de datos comunes y se armonizan procesándolos con un conjunto común de canalizaciones, centralizaciones o unificaciones alrededor de un objetivo común para que se puedan integrar y analizar

²²Berlin Declaration on NFDI Cross-Cutting Topics.(2019). Disponible en: <https://user.fz-juelich.de/record/872778>



fácilmente. En este sentido, los datos comunes reducen el costo y el esfuerzo necesarios para el análisis significativo de los datos de investigación.

Los Datos Commons ahorran tiempo a los investigadores al integrar y admitir herramientas de software, aplicaciones y servicios de uso común. Los datos comunes utilizan diferentes estrategias para esto. Los Data Commons en sí pueden facilitar espacios de trabajo que respalden el análisis de datos. También ahorran dinero y recursos para una comunidad de investigación, ya que cada grupo de investigación de la comunidad no tiene que crear su entorno informático y alojar los mismos datos.

Las anteriores características son complementadas con las lecciones aprendidas de varios proyectos, por ejemplo, las características observadas por el Open Commons Consortium (Open Commons Consortium, 2020) y Coleridge Initiative (Coleridge Initiative, 2020):

- Los conjuntos de datos en los bienes comunes deben tener identificaciones digitales estables y persistentes que incluyan control de acceso a los datos y metadatos y puedan devolver los metadatos.
- Una API, siglas en inglés de Interfaz de Programación de Aplicaciones, puede facilitar a las personas la búsqueda, el uso de datos y la autenticación de los usuarios que acceden a la información.
- La portabilidad es importante para permitir el intercambio y el transporte de datos entre los datos comunes.
- Los acuerdos de intercambio de datos entre dos proveedores de servicios de datos comunes son una forma eficaz de transferir datos sin costo para que un investigador pueda acceder a los datos almacenados en otros datos comunes.
- Asignar, restringir el tiempo o cobrar por el acceso a los recursos informáticos es una forma razonable de mitigar o descargar los costos de administrar un gran conjunto de datos.
- Proporcionar un entorno seguro dentro de las diferentes agencias gubernamentales (proveedores de datos) que pueden colocar y compartir sus datos.
- Crear un sistema eficiente y seguro para las personas, los datos y los proyectos que interactúan en los datos comunes.
- Crear programas de capacitación para utilizar los datos comunes. Los programas de capacitación para los empleados de las agencias gubernamentales (contribuyentes de datos) desarrollan las habilidades y los conocimientos necesarios para aprovechar al máximo la gran cantidad de datos disponibles.
- Contar con un proceso de desidentificación. Todos los microdatos que se agregan a los datos comunes tienen ciertos elementos de identificación de datos



personales que deben ser anulados antes de estar disponibles (por ejemplo, nombres, número de seguro social, identificadores comerciales).

- El proveedor de servicios de datos comunes tendrá diferentes herramientas que faciliten la comunicación efectiva entre equipos o usuarios del sistema, generando que se conformen comunidades de usuarios.

4.2.3 Tecnología

Los Data Commons funcionan mejor en situaciones en las que hay mucha confianza entre los actores implicados. Todos los participantes reciben beneficios basados en el recurso común compartido. Aunque los Data Commons pueden construirse utilizando la misma estructura tecnológica que los Data Trusts, en general la tecnología no tiene que ser tan restringida, y se puede poner menos enfoque en temas de verificación y auditoría de los datos. Los Data Commons tienen que contar con estándares de datos y APIs, pero pueden tener un entorno más parecido a una plataforma de datos abiertos o portales de datos en general.

En cuanto a la seguridad de la información, todos los datos que se compartan en una explotación en un Data Commons deben estar anonimizados. Asimismo se debe evitar el uso de información personal y casos específicos. Sin embargo, si se requieren casos específicos (como, por ejemplo, en historias clínicas médicas) solo el recolector de esos datos puede tener acceso a los datos que habiliten la identificación y estos deben estar encriptados. El resto de actores del modelo de datos sólo tendrán acceso a un ID de identificación del caso. En este contexto, se debe cumplir con la guía para la anonimización de bases de datos en el Sistema Estadístico Nacional (DANE, 2018).

En cuanto a seguridad, X-Road puede ser una alternativa útil para facilitar la verificación de los datos. No obstante, se debe poner especial atención a la protección de los datos personales. Esto debe involucrar a todos los actores del Data Commons (en especial a los que participan en las etapas de recolección y procesamiento) y debe responder a las buenas prácticas de ingeniería en software y hardware. Asimismo, y debe responder a la normativa actual de protección de datos personales, al Modelo de Seguridad y Privacidad de la Información publicado por MinTIC (MinTIC, 2016) y a los capítulos de seguridad dentro de las guías del Marco de Referencia de Arquitectura (MinTIC, 2019).

4.2.4 Flujos de datos

Los datos de un Data Commons pueden estar centralizados en un repositorio, como un lago de datos, o descentralizados. La mejor opción depende de las circunstancias específicas del proyecto de Data Commons. Un aspecto importante del flujo de datos en un Data Commons es que todos los datos se comparten por igual, al igual que los



derechos sobre los mismos. Esto ayuda a facilitar la igualdad de acceso a los datos por parte de todos los participantes en un Data Commons.

4.2.5 Oportunidades

El impacto futuro de un Data Commons podría incluir lo siguiente:

- **Aumento de la disponibilidad de los datos como recurso común.** Este mayor acceso podría conducir a otros beneficios como una mayor responsabilidad, una mayor innovación o mejoras en la toma de decisiones y la eficiencia.
- Organizaciones y productos digitales más fiables. Un beneficio de los Data Commons podría ser que las organizaciones y los productos **se ajusten más a las necesidades de sus interesados**. Esto conducirá a organizaciones y productos digitales más fiables.
- **Metodologías replicables** para crear bienes comunes de datos y, en general, enfoques no legales de la gobernanza de datos

4.2.6 Ejemplos colombianos

Aclímate

[Aclímate](#) es una [colaboración intersectorial](#) de organizaciones públicas, privadas y no gubernamentales en Colombia. La colaboración está liderada por el Centro Internacional de Agricultura Tropical (CIAT) y el Ministerio de Agricultura de Colombia. La plataforma en línea Clima y Sector Agropecuario Colombiano (CSAC) ofrece datos meteorológicos útiles para los agricultores. Estos datos incluyen predicciones climáticas y datos históricos sobre precipitaciones, radiación solar y temperaturas máximas y mínimas. El sitio también ofrece datos agronómicos sobre el cultivo de arroz y maíz, entre otros datos.

Este Data Commons ayuda a los agricultores colombianos a tomar decisiones sobre el mejor momento para plantar y cuáles variedades de cultivos crecerán mejor en las condiciones previstas. Este Data Commons también ha tenido un impacto económico demostrado. En 2013, por ejemplo, CSAC pronosticó una sequía que provocaría la pérdida de muchos cultivos. Los agricultores que siguieron las recomendaciones del CSAC ahorraron USD \$3.6 millones en insumos agrícolas que se habrían gastado en la cosecha fallida.



Improving Food Security and Agricultural Competitiveness (IFSAC)

[IFSAC](#) fue una iniciativa de compartir imágenes satelitales con el propósito de contribuir al sector agrícola en los países de América Latina, incluida Colombia. Satellite Applications Catapult trabajó con las partes interesadas del sector agrícola para educar y demostrar la utilidad de las tecnologías satelitales para optimizar la producción de arroz, soja y uva. El proyecto fue financiado por el Foreign and Commonwealth Office (FCO) del Reino Unido. Las embajadas del Reino Unido en Perú, Colombia y Paraguay ayudaron en el desarrollo del proyecto.

Este Data Commons ayudó a crear varios productos innovadores que ayudan a optimizar la producción de cultivos mediante la mejora de las decisiones de los productores. Los datos recopilados se usaron para monitorear las condiciones y el desarrollo de los cultivos, identificar dónde se estresaron los cultivos para ayudar en las intervenciones y monitorear un fungicida piloto para cuantificar la efectividad de los diferentes métodos de tratamiento y optimizar el control de las enfermedades de las plantas. En Colombia específicamente, este Data Commons se centró en el cultivo de arroz y recopiló datos satelitales sobre este subsector.

4.2.7 Ejemplo internacional

Barcelona Ciudad Digital

La ciudad de Barcelona ha creado un Data Commons centrado en los datos como recurso común. La visión de [Barcelona Ciudad Digital](#) es democratizar el acceso a los datos otorgando a los residentes derechos sobre sus propios datos. Su intención es empoderar a los residentes de Barcelona para "avanzar hacia una sociedad digital en la que las personas puedan controlar, compartir y explotar los datos" (Data Commons Barcelona, 2019).



Género y la movilidad urbana en Chile

UNICEF, la Universidad de Desarrollo, Telefónica, the ISI Foundation, y Digital Globe [establecieron un data commons para explorar las brechas de género](#) en la movilidad urbana en Santiago de Chile. Más del 90% de la población de Chile vive en áreas urbanas y el 40% de la población del país vive en la ciudad de Santiago. Santiago se está expandiendo, lo que pone más presión sobre las poblaciones vulnerables que tienen que viajar distancias más largas para vivir, estudiar y trabajar. Más mujeres que hombres en la ciudad viajan a pie y utilizan el transporte público: el 40% de las mujeres en Santiago utilizan el transporte público con regularidad, frente al 30% de los hombres. El transporte público, por lo tanto, debe diseñarse teniendo en cuenta las necesidades de las mujeres que lo utilizan con regularidad. Mediante el uso de datos compartidos de Call Detail Records (CDR), con datos del censo, datos de transporte público en el formato General Transit Feed Specification y OpenStreetMap, estas organizaciones trabajaron para comprender mejor cómo el género influye en las trayectorias de movilidad y explorar las brechas de género en la movilidad urbana. Los participantes en el Data Commons compartieron datos entre sí.

Este Data Commons ayudó a las organizaciones a evaluar la accesibilidad de los servicios de transporte en Chile y ver las formas en que se podría mejorar el transporte público para servir mejor a las mujeres en Santiago.

4.3 Data Marketplace

Los Data Marketplaces o Mercados de Datos se definen como una plataforma digital para el intercambio de activos de datos entre personas y empresas actuando como consumidores y proveedores, en donde se crean catálogos de datos, conforme con los casos de uso dirigidos al sector empresarial, al sector público, sectores de nicho o al público en general. Estas plataformas de intercambio de datos deben actuar como un intermediario neutral, y permitir que cualquier usuario cargue e intercambie sus conjuntos de datos.

Los mercados de datos se diferencian de los Data Trusts y los Data Commons en su modelo de gobernanza. En un mercado de datos, hay un organismo, que puede ser público o privado, que define cuáles actores pueden participar en el mercado. Además, en muchos casos hay un regulador que supervisa que las transacciones en el mercado de datos sean justas y fija los precios de los datos.

Los mercados de datos tratan de estructurar y hacer más eficiente el mercado de datos existente, por ejemplo, reuniendo la actividad en un único portal digital en línea, utilizando un enfoque coherente para la valoración de los datos o aplicando las normas del mercado.

Aunque la mayor parte de la actividad actual se desarrolla en el sector privado, **gobiernos locales y nacionales están explorando el papel que desempeñan en la provisión de componentes técnicos o no técnicos a los mercados de datos que regulan.**

En un primer nivel conceptual tenemos el esquema para conceptualizar el mecanismo de intercambio de datos dentro de un mercado de datos:

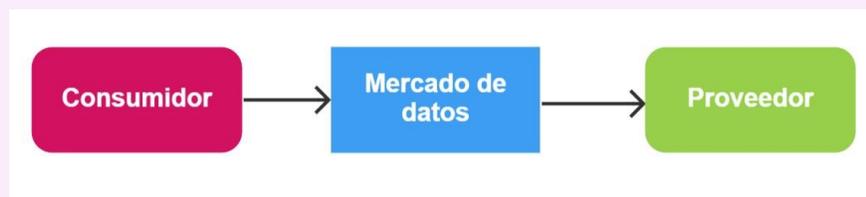
Figura 12: Primer mecanismo de intercambio de datos en un mercado de datos



Fuente: Elaboración propia

Este esquema muestra el flujo de los datos, con una entidad central que se encarga de la administración de los datos. Es importante aclarar que el Data Marketplace solo administra y por lo tanto se puede limitar a almacenar los metadatos y no los datos como tal. En la otra dirección está el incentivo económico para el proveedor de los datos.

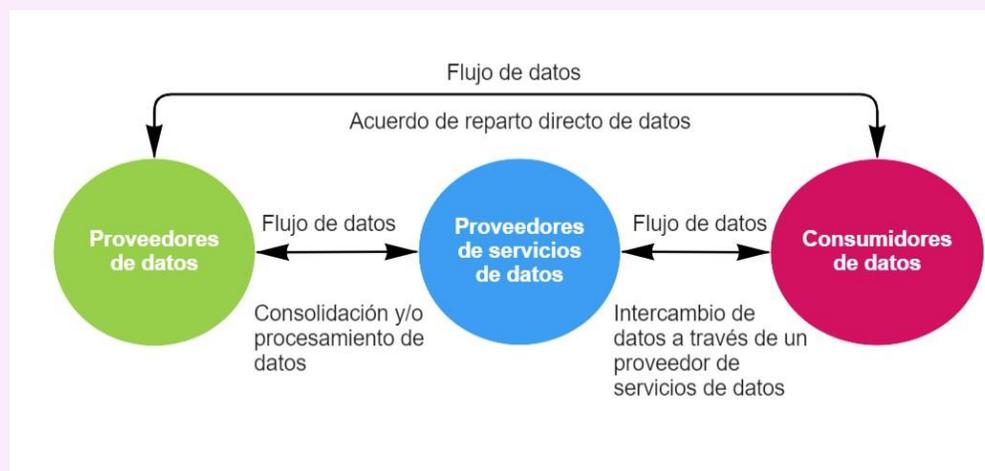
Figura 13: Segundo mecanismo de intercambio de datos en un mercado de datos



Fuente: Elaboración propia

En esta forma bidireccional puede haber un incentivo (tarifa) para justificar o sustentar el Mercado de Datos.

Figura 14: Flujo de datos en un mercado de datos



Fuente: Elaboración propia basada en IMDA (2019) (Infocomm Media Development Authority, Personal Data Protection commission Singapore, & Intellectual Property Office of Singapore, 2019)

Tras una mayor exploración de la naturaleza de las tres entidades que interactúan en el nivel uno, podemos identificar que cada una puede tener diferentes naturalezas de la siguiente manera: el proveedor puede ser público o privado, una persona o una Institución.

4.3.1. Definiciones

A continuación un listado de definiciones de Data Marketplace:

Figura 15: Definiciones de Data Marketplace

Organización	Definición
Data Economy Lab ²³	Permitir la monetización de datos a través de una plataforma en la que el administrador permite tanto el descubrimiento como las transacciones entre los compradores y los generadores de datos.
Mozilla ²⁴	Un enfoque que permite a los consumidores vender o intercambiar sus datos personales por servicios u otros beneficios.

²³ The Data Economy Lab. (2022). Data Stewardship models. <http://tool.thedataeconomylab.com/our-data-models>

²⁴ Mozilla. (2020). What Does it Mean? | Shifting Power Through Data Governance. Disponible en: <https://foundation.mozilla.org/en/data-futures-lab/data-for-empowerment/shifting-power-through-data-governance/#what-is-a-data-marketplace>



Snowflake ²⁵	Un mercado de datos es un lugar o almacén transaccional en línea que facilita la compra y venta de datos.
AI Multiple ²⁶	Es una plataforma en la que los usuarios compran o venden diferentes tipos de conjuntos de datos y flujos de datos de varias fuentes.
Nokia ²⁷	El valor de un mercado de datos es facilitar el proceso de comercio e intercambio de datos de la manera más eficiente y segura posible, eliminando al mismo tiempo los obstáculos que impiden a los compradores y vendedores realizar sus actividades.
Shanghai Municipal People's Congress ²⁸	Permiten que los datos se clasifiquen, coticen y negocien como materias primas.
Anjana Data ²⁹	Un gran bazar controlado en el que todos los activos de información son puestos a disposición de los consumidores potenciales por diferentes productores dentro del entorno gobernado de una organización.
A-Team Insight ³⁰	Poner en contacto a los proveedores de datos con los consumidores a través de una plataforma central, normalmente basada en la nube.
Dawex ³¹	Donde las organizaciones pueden distribuir, obtener, intercambiar, compartir y comercializar datos y/o orquestar ecosistemas de datos, creando poderosas asociaciones de datos.

Fuente: Elaboración propia.

4.3.2. Data Marketplace para el sector público

²⁵ Snowflake. (2022). What is a Data Marketplace?. Disponible en: <https://www.snowflake.com/guides/what-data-marketplace>

²⁶ Cem Dilmegami. (2021). Data Marketplaces: What, Why, How, Types, Benefits, Vendors. Disponible en: <https://research.aimultiple.com/data-marketplace/>

²⁷ Nokia. (2022). Data Marketplace. Disponible en: <https://www.nokia.com/networks/services/nokia-data-marketplace/>

²⁸ Cooley. (2022). PRC's New Efforts to Facilitate Data Trading: Shanghai Data Exchange Kicks Off Trading. Disponible en: <https://cdp.cooley.com/prcs-new-efforts-facilitate-data-trading-shanghai-data-exchange-kicks-off-trading/>

²⁹ Anjana Data. (2021). How to successfully build an internal Data Marketplace for your organization. Disponible en: <https://www.linkedin.com/pulse/how-successfully-build-internal-data-marketplace-your-mario/>

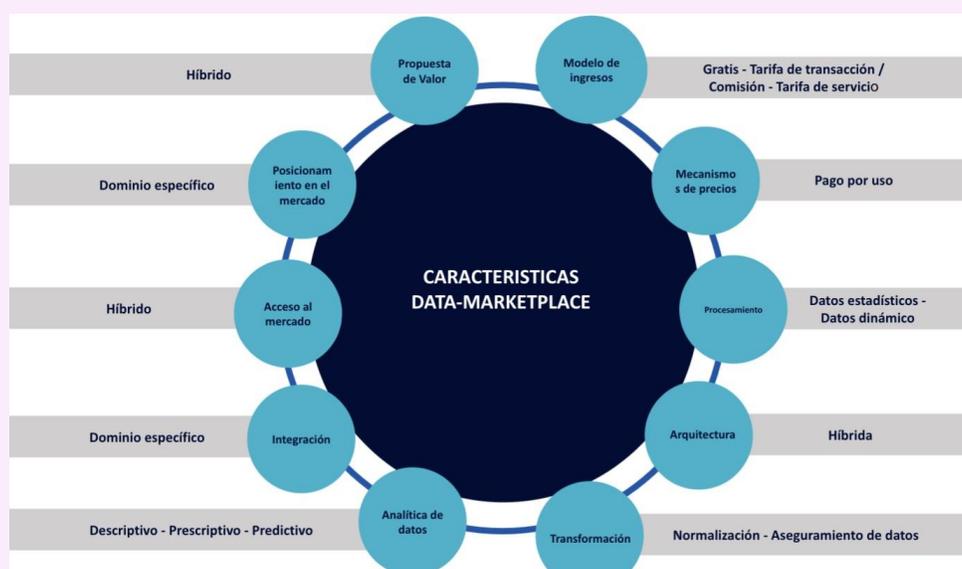
³⁰ A-Team Insight. (2021). The Rise and Rise of the Financial Data Marketplace. Disponible en: <https://a-teaminsight.com/the-rise-and-rise-of-the-financial-data-marketplace/>

³¹ Dawex. (2022). Data Exchange Platform. Disponible en: <https://www.dawex.com/en/data-exchange-platform/>

En aquellos casos en los que el marco legal de un país prohíbe la participación del sector público en mercados de datos como proveedor, existen otras formas en las que este se puede beneficiar de la implementación de este modelo de gobernanza de intercambio de datos. A continuación presentamos una lista de las oportunidades que presenta un **Data Marketplace para el sector público**:

1. El sector público puede ser partícipe de un mercado de datos validando que las transacciones cumplen con el marco regulatorio y la ley de protección de datos nacional.
2. Si el gobierno es el que gestiona la plataforma del mercado de datos, a través del monitoreo de las transacciones y actividades que se realizan dentro de éste, puede obtener visibilidad y conocimientos del sector donde se ha establecido el mercado de datos.
3. Las entidades públicas no pueden comercializar sus datos, pero pueden proveerlos en otros contextos. Sin embargo, el sector público puede ser un consumidor de datos en los mercados de datos. A través de las prácticas de adquisición de datos, puede fomentar la creación de nuevos datos y exigir normas específicas a los proveedores.
4. Los gobiernos tienen la oportunidad de establecer sistemas de valoración de los datos que se puedan adoptar a nivel nacional.
5. El sector público puede proporcionar soporte y apoyo en cuanto a habilidades técnicas para la implementación y gestión de mercados de datos.
6. El sector público puede crear metodologías replicables de diseño e implementación de mercados de datos para que otros actores puedan reutilizarlas.

Figura 16: Características de un mercado de datos



Fuente: C4RI, 2021.



Adicionalmente, el mercado de datos debe considerarse un sistema transaccional, que solo almacena metadatos, de modo que los datos estén siempre en los servidores de los proveedores y se transfieran directamente a los consumidores con la aprobación y supervisión del mercado de datos. Aquí es donde se establecen los términos de las negociaciones, se crean los acuerdos, se autoriza el intercambio de datos y se toma el registro histórico de las transacciones de compra-venta (C4IR, 2021).

El acuerdo de compra-venta puede incluir lo siguiente:

- Identidad del comprador (certificado emitido por la capa de interoperabilidad)
- Identidad del proveedor (certificado emitido por la capa de interoperabilidad)
- Metadatos: tamaño de los datos, tiempo de actividad, número de descargas permitidas, entre otros.
- Firma de datos única: hash, para verificar que los datos que se entregan sean los mismos que los del proveedor de registro.
- Costo de los datos (certificado emitido por la capa del modelo financiero)
- Impuestos del comprador y vendedor
- Comisiones del mercado de datos
- Aceptación de los términos y condiciones generales del mercado de datos (firmados en el momento de la compra)
- Información detallada o categorías, sobre el uso que se le dará a los datos (entregados en el momento de la compra)

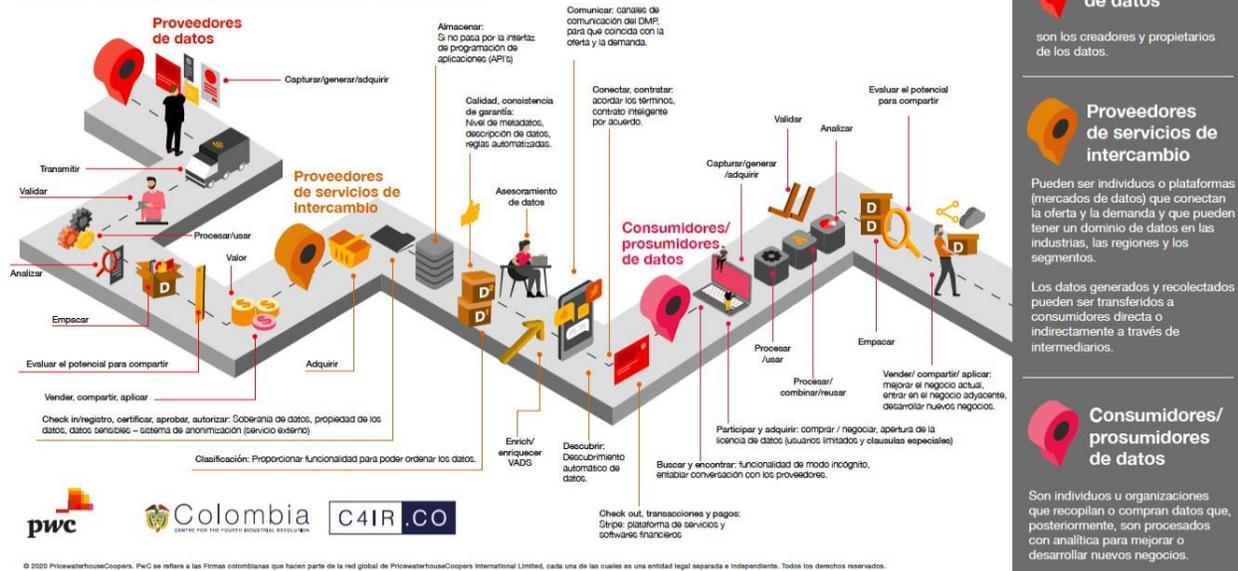
Este acuerdo de compraventa puede conducir a un contrato inteligente, que ambos actores firman y su ejecución da como resultado:

- Token de acceso a datos
- Marca de tiempo (hora en que se realizaron las transacciones)
- Hash de la transacción
- Número de bloque
- Certificado DATAMARKET
- Documento con el contrato firmado entre las partes.

Figura 17: Conoce el camino de los mercados emergentes de datos

Conoce el camino de los mercados emergentes de datos

Autoría de PwC, bajo el Proyecto Moonshot liderado por el C4IR.CO



Fuente: PwC, 2020.

En lo que respecta a los datos, son múltiples los atributos que determinan el valor de los mismos. A continuación, se presentan aquellos que se han definido como más relevantes dentro del Proyecto Moonshot, con el objetivo de ayudar a las partes interesadas a alinearse en la valoración de los datos (WEF, 2021).

Figura 18: Atributos de valor de los datos

Integridad Por lo general, cuanto más completo sea un conjunto de datos, más valioso es. La integridad de los datos tiene que ser determinada en función a su caso de uso.	Coherencia Los datos son coherentes si se ajustan a la sintaxis de su definición.	Precisión El grado en el que los datos describen correctamente el "mundo real". La procedencia es un aspecto crítico de la precisión, ya que permite al usuario comprender la historia de los datos.	Oportunidad Cuanto más actualizados estén los datos, más valiosos serán.
Exclusividad Cuanto más único el conjunto de datos, más valiosos son los datos.	Restricciones de uso Cuanto menos restringido el uso de los datos compartidos, mayor es su valor.	Responsabilidad y riesgos Los riesgos y responsabilidades asociados a los datos compartidos podrían reducir su valor.	Interoperabilidad Un consumidor elegirá el conjunto de datos más accesible. La interoperabilidad de los datos suele ser crítica para que el consumidor pueda utilizar los datos.



Utilidad

El valor de los datos reside en su **utilidad**.
Los datos sin una aplicación concreta o **aplicación** concreta o potencial no tiene valor inherente.
Los **impulsores del valor** de los datos son los factores clave subyacentes que conforman la utilidad de los datos.

Fuente: IMDA Singapore, Key Data Value Drivers - Guide to Data Valuation for data sharing, 2019



4.3.2. Capas de la gobernanza

En un Data Marketplace, el modelo de gobernanza además de establecer la forma de interacción entre las partes, también establece el modelo financiero, que debe contener (C4RI, 2021):

- El modelo de pago por datos
- El modelo de ventas de datos
- El modelo de financiación del mercado de datos.

Para ello, una de las primeras tareas a realizar es el establecimiento del modelo de valoración de datos. Existen modelos de valorización de datos que tienen en cuenta múltiples factores. Estos se pueden definir inicialmente mediante las siguientes relaciones:

- Costo-Consumo
- Rentabilidad de costes
- Consumo de calidad
- Oferta-demanda

Adicionalmente, el modelo de gobernanza del mercado de datos debe dictar su modelo de pagos, de manera que existan incentivos para los tres agentes que interactúan en el sistema:

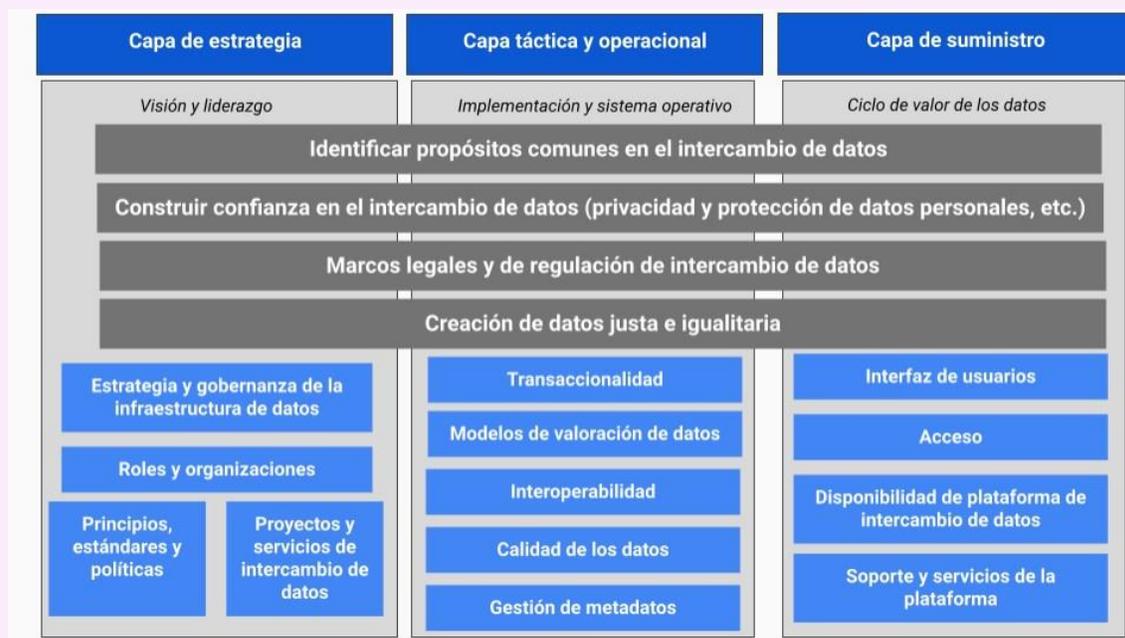
- Proveedor de datos
- Consumidor de datos
- Proveedor de mercado de datos

Cabe señalar que además del valor intrínseco de los datos, se debe implementar el modelo de pagos dictado por el gobierno (incluidos impuestos y pagos al mercado de datos), de modo que se pueda generar un incentivo suficiente para el comprador, el proveedor del mercado de datos y proveedores. Los pagos, en principio, se realizan en la moneda legal de cada país, es decir, no se evidencia un esquema de pago en criptomonedas (C4RI, 2021). Por tanto, se puede descartar el uso de criptomonedas como modelo de pago.

Con respecto al momento de la oferta de compra, se debe declarar el propósito para el cual se utilizarán los datos, de modo que las partes involucradas en la transacción puedan compartir una visión clara y transparente del uso permitido de los datos, en cumplimiento de normativas y requisitos contractuales.

A continuación, se presenta un ejemplo de modelo de gobernanza para un Data Marketplace, en este caso el propuesto por el proyecto Moonshot del World Economic Forum y el Centro para la Cuarta Revolución Industrial (WEF, 2021).

Figura 19: Modelo de Gobernanza para un Data Marketplace



Fuente: World Economic Forum, 2021

4.3.3 Tecnología

En un Data Marketplace, el modelo de interoperabilidad del sistema puede ser centralizado o descentralizado (C4RI, 2021). En los modelos centralizados, hay una entidad que emite certificados para el intercambio seguro de datos. X-Road y Blockchain privada son ejemplos de esto. En los modelos descentralizados, el intercambio seguro de datos se realiza mediante un acuerdo entre todas las partes. Ejemplos de esto son entidades federadas y blockchain públicas. Esto no implica ni descarta, per se, el uso de la tecnología Blockchain.

Además, el modelo operativo debe estar basado en flujos de trabajo automatizables para la aplicación de políticas y procedimientos. Para ello, es necesario considerar las siguientes tecnologías para su correcto funcionamiento (de Francisco Ruiz, 2021):

- Tecnologías para el movimiento e ingestión de datos (ETL/ELT, FTP, eventos, streaming, ...)



- Tecnologías para el almacenamiento de datos (bases de datos relacionales, NoSQL, DWH, Data Lake, ...)
- Tecnologías para el procesamiento de datos (preparación, transformación, modelado, calidad, ...)
- Tecnologías para el consumo y explotación de datos (BI, Reporting, Analytics, Self-Service, ...)
- Tecnologías para la integración de datos (ETL/ELT, virtualización, APIs, DaaS, ...)

Finalmente, todo el sistema operativo requiere diferentes interfaces (front-end) para funcionar. Se puede establecer una lista inicial a partir del tipo de usuario de Data Market (C4RI, 2021):

- Administrador
- Comprador
- Vendedor
- Auditor
- Público en general

Cada uno de estos usuarios del sistema requiere una interfaz dedicada con funcionalidades extremadamente específicas. Debido a que un usuario de Data Marketplace puede tener múltiples roles (por ejemplo, un comprador también puede ser un vendedor), esas interfaces dedicadas deben ser consistentes entre sí (C4RI, 2021).

4.3.4 Flujos de datos

Al igual que en los Data Trusts y los Data Commons, los datos en un Data Marketplace pueden estar centralizados o descentralizados. El aspecto más importante que tiene que ver con los flujos de datos dentro de un Mercado de Datos tiene que ver con el flujo financiero dentro de la plataforma. El dinero siempre fluye dentro del mercado de datos, y un mercado de datos puede ser construido de manera que los creadores puedan ver y regular las transacciones y los precios de los datos dentro del mercado. El flujo de datos y dinero en el mercado, sin embargo, siempre cambiará dependiendo del contexto y el propósito del mercado.

4.3.5 Oportunidades

El impacto futuro de un data marketplace podría incluir lo siguiente: **Crearé nuevos ingresos para las personas.** Los Data Marketplace permiten hacer participar a las personas del valor de los datos, con lo cual, obtienen un beneficio adicional que puede ser económico u otro tipo de bien o servicio.

- Permitirá comprender mejor las distintas **funciones que puede desempeñar la administración pública en los mercados de datos.** Algunos ejemplos de lo anterior son: venta de datos, compra de datos, fijar de precios, definir una



metodología de valoración común o hacer cumplir las normas relativas a las transacciones.

- Conducirá a una **mayor comprensión** de cómo la infraestructura técnica y las metodologías digitales pueden apoyar el intercambio de datos en general, y los mercados de datos en particular.
- A largo plazo, el aumento de la capacidad de descubrimiento de datos a través de los mercados de datos conduce a un **aumento del valor económico directo causado por la compra/venta de datos** que puede tener diferentes implicaciones económicas para el valor creado de los datos. Las variaciones en los precios (sujetos al mercado) pueden tener un impacto tanto positivo como negativo en términos de acceso y el uso de conjuntos de datos.
- Los gobiernos podrán **desarrollar metodologías generales para la creación y funcionamiento de mercados de datos**; así como un marco de política pública para decidir cuándo y cómo intervendrá el gobierno en los mercados de datos. Estas metodologías y marco pueden utilizarse para estimular y dar forma al mercado de datos de un país.
- Los operadores también tendrán **un mayor control sobre el mercado**, por ejemplo, al tener la capacidad de fijar un precio o determinar cómo los conjuntos de datos son descubiertos por las personas que buscan datos. Se trata de efectos similares a los observados en los mercados digitales, donde el papel de organizaciones como Amazon y eBay es cada vez más reconocido. Gobiernos como Colombia deben tener en cuenta lo anterior al diseñar políticas de datos y definir el grado de intervención pública en el mercado.

4.3.6 Ejemplo internacional

BitsaboutMe

[BitsaboutMe](#) ayuda a los individuos a almacenar y vender sus datos personales. BitsaboutMe cuenta con una plataforma segura en la que los particulares pueden recoger y almacenar sus datos. La plataforma les proporciona información sobre sus compras y su estilo de vida. Los individuos también pueden optar por vender sus datos anónimos. Pueden ganar dinero o reinvertir lo que ganan en iniciativas de sostenibilidad medioambiental.

Las empresas que compran datos en BitsaboutMe reciben datos anónimos de alta calidad procedentes de diversas fuentes que cumplen las leyes de privacidad, lo que ayuda a las empresas a acceder a datos éticos de alta calidad sobre sus mercados objetivo. Sólo las empresas pagan por utilizar BitsaboutMe. Para los individuos, una cuenta en la plataforma es gratuita y no hay obligación de compartir nunca los datos. BitsaboutMe es un ejemplo de cómo un mercado de datos puede facilitar el uso, la recogida y la venta de datos de forma ética y transparente.

4.4 Entidades líderes

Existen diferentes entidades dentro del gobierno colombiano que deberían liderar proyectos piloto de los tres modelos de gobernanza de datos. El **Departamento Administrativo Nacional de Estadísticas (DANE)**, entidad responsable de la planificación, recopilación, procesamiento, análisis y difusión de las estadísticas oficiales de Colombia, debería participar en cualquier proyecto piloto de intercambio de datos.

En Colombia, los proyectos piloto de Data Trusts podrían ser liderados por aquellos **ministerios que coordinan las políticas del sector al que pertenece el caso de uso abordado**. Por ejemplo, si se quiere explorar un caso de uso relacionado con movilidad urbana, sería pertinente involucrar al MinTransporte o si se busca trabajar en cuestiones de salud habría que incluir al MinSalud. Asimismo, será fundamental contar con la contribución de la **Superintendencia de Industria y Comercio** en cuestiones de protección de datos personales.

Asimismo, cualquier proyecto piloto de Data Commons en el que la **investigación sea el foco principal**, debe incluir al **Ministerio de Ciencias, Tecnología e Innovación (MinCiencias)** como entidad líder del proyecto. Los proyectos piloto de Data Commons que se centran en mejorar el acceso a los datos como recurso común pueden ser liderados por los ministerios sectoriales que consideren que este enfoque podría ayudar a mejorar el intercambio de datos dentro de su sector.



El Ministerio de Comercio, Industria y Turismo (MINCIT) debería liderar cualquier proyecto piloto de Data Marketplace, debido a su misión de apoyo a las actividades empresariales, a los productores de bienes, servicios y tecnología, así como a la mejora de la competitividad y al fomento de la generación de mayor valor añadido en el país.

5. Recomendaciones normativas e institucionales

5.1 Recomendaciones de política

En varios países se está trabajando para mejorar la política y la legislación en torno a la gobernanza de datos, la protección de datos y el intercambio de datos, a menudo centrándose en los algoritmos y el uso responsable.

Con la aprobación de la legislación sobre el [GDPR \(General Data Protection Regulation\)](#) en 2016, la Unión Europea estableció su legislación sobre protección de datos, modificando los requisitos de responsabilidad, seguridad de los datos, protección de los datos por defecto, consentimiento y derechos de privacidad personal, entre otras áreas. La Ley de Datos de la UE también propuso nuevas medidas para la gobernanza de los datos, con una base de protección de datos. Asimismo, a nivel regional, la [Ley General de Protección de Datos de Brasil](#) ha abordado el tema de forma innovadora. **Dichos lineamientos están en constante revisión y actualización, de acuerdo a las necesidades de sus ciudadanos y el contexto global.**

Con estas nuevas políticas, se están **introduciendo nuevos estándares para la protección de datos y la privacidad**, específicamente en lo que respecta a la privacidad por defecto. Se **recomienda explorar las últimas actualizaciones a estas nuevas piezas legislativas como referencia de mejores prácticas a nivel internacional**. Es importante explorar modelos similares al Plan Nacional de Infraestructura de Datos y el Modelo de Gobernanza de Infraestructura de Datos Colombiano y analizar de qué manera estas ideas podrían adaptarse al contexto colombiano.

Dado el contexto rápidamente cambiante que rodea a la protección de datos, el intercambio de datos, la gobernanza de datos y los algoritmos, hay que trabajar para **determinar qué instituciones trabajarán para desarrollar guías y mejores prácticas asociadas a los datos, el intercambio de datos y la gobernanza de datos.**

El Open Data Institute (ODI) tiene un término para describir el **flujo confiable de datos dentro de los sectores o ecosistemas**. Definen el término "data assurance" como "el proceso o conjunto de procesos que aumentan la confianza en que los datos satisfacen



una necesidad específica, y que las organizaciones que recopilan, acceden, utilizan y comparten datos lo hacen de forma fiable". El ODI ha publicado un [informe sobre la forma en que varios tipos de organismos en el Reino Unido](#), incluidos los reguladores, los organismos industriales y los organismos profesionales, pueden ayudar a crear confianza en los datos y en el intercambio de datos en varios sectores (ODI, 2022). **Se recomienda que Colombia haga un análisis similar de los diferentes sectores en su país y ecosistema para determinar las mejores formas en que las diversas instituciones puedan ayudar a construir la confianza en los datos y el intercambio de datos.**

5.2 Recomendaciones para los reguladores

Por último, en cuanto a las recomendaciones para los organismos que establecen y aplican normas y reglamentos, se recomienda que los reguladores de múltiples sectores **evalúen su capacidad en torno a los datos y su intercambio**. ¿Están desarrollando los reguladores de sectores específicos la capacidad de regular las actividades de datos o el uso de datos como parte de sus actividades?, ¿Son los reguladores conscientes de los flujos de datos o tienen la capacidad de examinarlos? **Se recomienda que Colombia considere el desarrollo de un proyecto de investigación específico para determinar la capacidad de los reguladores para participar en el intercambio de datos de una manera ética y confiable.**

6. Conclusiones

A continuación se presentan algunas de las conclusiones a las que se llegó en la presente investigación sobre la gestión de datos desde escenarios como los de Data Commons, Data Trust y Data Marketplace.

En la actualidad existe una fuerte erosión de la confianza pública y la confianza en las organizaciones de recopilación de datos y en las tecnologías que se basan en estos datos que debe ser intervenida desde el aparato estatal, a fin de que la sociedad pueda acceder y utilizar datos confiables para el bien público. Esta necesidad apremiante de explorar nuevos modelos de gobernanza de datos que brinden a las personas cierto control sobre sus datos personales debe ser ejercida desde el gobierno nacional a través de nuevas dinámicas legislativas que les ayuden a mirar hacia adelante y a gestionar los posibles riesgos de este tipo de operaciones.

Como se mencionó a lo largo de todo este documento, tres de las nuevas formas de gestión de información son Data Trust, Data Commons, y Data Marketplace. Si bien aún no se ha llegado a definiciones definitivas sobre estas figuras, a partir de los estudios adelantados por expertos mundiales, sí tenemos claro que son modelos de intercambio



de datos que vale la pena explorar. Aunque a primera vista parecen similares y tienen puntos en común, se diferencian en que un Data Commons presenta el objetivo de gestionar los datos como un bien común y por lo tanto es un modelo generalmente público, a diferencia de un Data Trust que, al llevar implícita la posibilidad de explotación, puede ser usado por el sector privado. Data Marketplace se diferencia de los dos en que asigna un valor monetario a los datos, por lo que puede ser utilizado más por el sector privado que por el público.

En ese orden de ideas, desde el Data Trust se presentan una serie de beneficios independientes, como crear nuevos negocios, ayudar a investigar una patología específica o capacitar a una comunidad de trabajadores, consumidores o ciudadanos, entre otras, generando una serie de rendimientos ya sea monetarios o de intercambio de bienes y servicios, pero siempre generando réditos.

Es por esto por lo que resulta tan importante que, **al momento de definir qué modelo es más conveniente para la gestión de los datos, se revise detenidamente cuál es el propósito de esa gestión.** Asimismo, hay que evaluar **cuáles son los objetivos que se pretenden abarcar por las partes interesadas en implementar un modelo de intercambio de datos.** Habiendo definido dichos puntos se puede decidir si el alcance de uno u otro modelo es suficiente para la gestión de la información.

Es imperativo que estas nuevas formas de gestión de datos sean aplicadas en Colombia, teniendo en cuenta que el universo tecnológico avanza a pasos acelerados y estamos en una carrera en donde el tiempo no juega a favor del propietario del dato, sino de quien pueda demostrar mayor efectividad en sus ejercicios de gobierno de datos.

El desafío para los titulares de datos al elegir un modelo adecuado de acceso a los datos es identificar la combinación correcta de mecanismos de gobernanza, supervisión y aplicación, procesos de revisión ética, mecanismos técnicos, estructuras legales y estrategias de participación de las partes interesadas para que todos puedan beneficiarse de las decisiones basadas en datos, verificando siempre que la seguridad no se vea vulnerada según el modelo que se seleccione de cara al usuario y a la protección de su información privada o sensible.

En este momento es clave establecer las reglas que le darán forma al aprovechamiento de los datos para el bien común. Para ello, hay que tener en cuenta las preocupaciones legítimas sobre la protección y el intercambio de datos y lograr que estas normativas vayan mucho más allá de garantizar derechos fundamentales.

Finalmente, hay que destacar que resulta relevante convocar a la mayor cantidad de sectores, actores e iniciativas, generando en ellos confianza en la forma en que se



gestiona tanto su información como sus necesidades. Solo de esta forma se puede garantizar modelos de gestión robustos, enriquecidos y realmente provechosos. La presencia activa de toda la sociedad hace que estas dinámicas de intercambio crezcan y produzcan mejores resultados de gestión, basados siempre en la transparencia y la confiabilidad de la información, así como en los esquemas de seguridad que la protegen.



7. Bibliografía

Aapti Institute. (2020). Data sharing for good in business and governance: Principles behind the process. Disponible en: <https://aapti.medium.com/data-sharing-for-good-in-business-and-governance-principles-behind-the-process-7f7b853c99ab>

Berlin Declaration on NFDI Cross-Cutting Topics.(2019). Disponible en: <https://user.fz-juelich.de/record/872778>

Brighthive. (2022). The Use of Data Trusts to Improve Impact (Responsibly). Disponible en: <https://brighthive.io/the-use-of-data-trusts-to-improve-impact-responsibly/>

Centre for Data Ethics and Innovation. (2021). Disponible en <https://www.gov.uk/government/publications/the-roadmap-to-an-effective-ai-assurance-ecosystem/the-roadmap-to-an-effective-ai-assurance-ecosystem>

Centro para la Cuarta Revolución industrial. (2021). Lineamientos Data Marketplace.

CIGI. (2019). Reclaiming Data Trusts. Disponible en: <https://www.cigionline.org/articles/reclaiming-data-trusts/>

Coleridge Initiative. (2020). *Coleridge Initiative*. Obtenido de <https://coleridgeinitiative.org/>

Common Wealth. (2020). A Common Platform:Reimagining Data and Platforms. Disponible en: <https://www.common-wealth.co.uk/reports/common-platform-tech-utility-antitrust#chapter-7>

Data Trusts UK. (2022). The Data Trusts Initiative. Disponible en: <https://datatrusts.uk/about>

DANE. (2018). *Guía para la anonimización de bases de datos*. Disponible en <https://www.dane.gov.co/files/sen/registros-administrativos/guia-metadatos.pdf>

DANE. (2020). *Norma Técnica de la Calidad del Proceso Estadístico*. Disponible en <https://www.dane.gov.co/files/sen/normatividad/NTC-Proceso-Estadistico-PE-1000-2020.pdf>

Data Commons Barcelona (2019). Our Vision. Disponible en: <https://datacommons.barcelona/our-vision/>

DWP Digital. (2022). DWP Digital is improving cross-government data sharing. Disponible en: <https://dwpdigital.blog.gov.uk/2022/02/17/dwp-digital-is-improving-cross-government-data-sharing/>



Elinor Ostrom. (1990). *Governing the commons. The Evolution of Institutions for Collective Action*. Disponible en https://wtf.tw/ref/ostrom_1990.pdf

Future Cities Canada. (2019). *Data Trusts: A New Tool for Data Governance*. Disponible en <https://portal.futurecitiescanada.ca/resources/data-trusts-a-new-tool-for-data-governance/>

GO FAIR. (2016). *FAIR Principles*. Disponible en <https://www.go-fair.org/fair-principles/>

IMDA Singapore. (2019). *Key Data Value Drivers - Guide to Data Valuation for data sharing*. Disponible en <https://www.imda.gov.sg/-/media/Imda/Files/Programme/AI-Data-Innovation/Guide-to-Data-Valuation-for-Data-Sharing.pdf>

Mills, S. (2019). *Who Owns the Future? Data Trusts, Data Commons, and the Future of Data Ownership*. Obtenido de <http://www.9yyxx.com/media/mmuacuk/content/documents/business-school/future-economies/Mills-2020.pdf>

MinTIC. (2016). *Modelo de Seguridad*. Disponible en <https://www.mintic.gov.co/gestion-ti/Seguridad-TI/Modelo-de-Seguridad/>

MinTIC. (2019). *Lenguaje Común de Intercambio de Información*. Obtenido de <http://lenguaje.mintic.gov.co/>

Modelo de Gobernanza de la Infraestructura de Datos Colombiano. (2022).

Mozilla. (2020). *What Does it Mean? | Shifting Power Through Data Governance*. Disponible en: <https://foundation.mozilla.org/en/data-futures-lab/data-for-empowerment/shifting-power-through-data-governance/#what-is-a-data-trust>

Mozilla Foundation (2022). *What Does it Mean? Shifting Power Through Data Governance*. Disponible en <https://foundation.mozilla.org/en/data-futures-lab/data-for-empowerment/shifting-power-through-data-governance/>

Nesta. (2022). *Unlocking the value in data as a commons*. Disponible en: <https://www.nesta.org.uk/feature/four-future-scenarios-personal-data-economy-2035/unlocking-the-value-in-data-as-a-commons/>

ODI. (2020). *Data Trusts in 2020*. Disponible en <https://theodi.org/article/data-trusts-in-2020/>

OECD.AI. (2021). *Enabling data sharing for social benefit through data trusts at the Global Partnership on AI, GPAI*. Disponible en: <https://oecd.ai/en/wonk/data-sharing-data-trusts>



Open Commons Consortium. (2020). *Open Commons Consortium*. Obtenido de <https://www.occ-data.org/>

Robert L. Grossman et. al. (2016) A Case for Data Commons: Towards Data Science as a Service. Disponible en <https://arxiv.org/abs/1604.02608>

Suzor, N., & Gray, J. E. (2018). *A data trust for industry data sharing*. Disponible en <https://digital.library.unt.edu/ark:/67531/metadc1616571/>

The Data Commons Blueprint (2020). Disponible en: <https://web.archive.org/web/20210119013750/https://datacommons.org.nz/>

The Data Economy Lab. (2022). Data Stewardship models. Disponible en: <http://tool.thedataeconomylab.com/our-data-models>

World Economic Forum. (2021). *Data for Common Purpose: Enabling Colombia's Transition to a Data-Driven Economy*. Disponible en <https://www.weforum.org/whitepapers/data-for-common-purpose-enabling-colombia-s-transition-to-a-data-driven-economy>

World Economic Forum. (2021). *Towards a Data Economy: An Enabling Framework*. Disponible en <https://www.weforum.org/whitepapers/towards-a-data-economy-an-enabling-framework>