

Datos abiertos más allá del sector público

PUBLICADORES, MOTIVACIONES Y MODELOS DE COLABORACIÓN



INTRODUCCIÓN	4
PARTE I: OTROS SECTORES PRODUCTORES DE DATOS	6
1.1. DATOS DESDE EL SECTOR PRIVADO	7
1.1.1 – Regulación en sectores específicos	8
1.1.2 – Transparencia institucional	9
1.1.3. Responsabilidad social corporativa	10
1.1.4 – Filantropía de los datos	12
1.1.5 – Servicios comerciales de intermediación de datos	13
1.2. - DATOS DEL SECTOR ACADÉMICO Y DE LA INNOVACIÓN	14
1.2.1 – Acceso abierto a los resultados científicos	15
1.2.2 – Datos de dominio público	17
1.2.3 – Laboratorios de experimentación con datos	18
1.2.4 – Sistemas colaborativos de datos	19
1.3 – LOS DATOS DEL TERCER SECTOR	20
1.3.1 - Transparencia en la gestión	21
1.3.2 - Datos para fomentar el desarrollo	22
1.3.3 – Datos para la colaboración multilateral	23
1.4 - LOS DATOS DE LAS PERSONAS	26
1.4.1 – Autogestión de los datos personales	26
1.4.2 – Repositorios de datos colaborativos	28
PARTE II - MODELOS DE COLABORACIÓN ENTRE ENTIDADES PÚBLICAS Y PRIVADAS	29
2.1 – ACUERDOS DE COMPARTICIÓN DE DATOS	30
2.2 – DATA TRUSTS	31

2.3 – DATA COLLABORATIVES	32
2.4 – DATA FELLOWSHIPS	33
2.5 – COOPERATIVAS Y COMUNIDADES DE DATOS	34
2.6 – RETOS, CONCURSOS Y COMPETICIONES	36
2.7 – INCUBADORAS Y ACELERADORAS	37
CONCLUSIONES	39
REFERENCIAS	41

Contenido elaborado por Carlos Iglesias, Open data Researcher y consultor

Este estudio ha sido desarrollado en el marco de la Iniciativa Aporta, desarrollada por el Ministerio de Asuntos Económicos y Transformación Digital, a través de la Entidad Pública Empresarial Red.es. Los contenidos y los puntos de vista reflejados en esta publicación son responsabilidad exclusiva de su autor. El equipo Aporta no garantiza la exactitud de los datos incluidos en el estudio. El uso de este documento implica la expresa y plena aceptación de las condiciones generales de reutilización referidas en el aviso legal que se muestra en:

<http://datos.gob.es/es/aviso-legal>

INTRODUCCIÓN

Desde sus orígenes el movimiento de los datos abiertos ha estado centrado en la apertura de los datos del sector público, e incluso la propia [declaración original de los principios de los datos abiertos](#) hace referencia única a los datos provenientes de las administraciones públicas (*open government data* - OGD). No obstante, esa percepción ha ido evolucionando y hoy en día es más frecuente hablar de los datos abiertos en un ámbito más general que incluye a todos los sectores, como demuestra por ejemplo la [declaración de los principios de los datos abiertos promovida más recientemente por el Open Data Charter](#), cuya formulación es perfectamente aplicable a cualquier sector.

Por otro lado, si bien en la actualidad contamos ya con una más que respetable y valiosa cantidad de datos disponibles en formato abierto provenientes de la administración, existen también algunas [limitaciones con respecto a los modelos de gobernanza y sostenibilidad de los datos públicos](#) que, junto a los retos pendientes de la transformación digital en el sector, están en cierto modo limitando la disponibilidad final de los datos. En este contexto, surgen también **otros sectores productores de datos** como fuentes alternativas o complementarias a los cientos de catálogos de datos abiertos que los gobiernos han venido publicando en los últimos años. Estos nuevos datos provienen de otros publicadores aparte del sector público – como el sector privado, el sector académico o el denominado tercer sector – todos ellos movidos por distintas motivaciones y por los posibles beneficios que veremos a lo largo del informe. Incluso la propia ciudadanía se ha convertido hoy en día en una amplia fuente de datos, aunque a veces sea de forma involuntaria.

Gracias a esta mayor variedad de productores de datos, se incrementan las posibilidades de enriquecer la información mediante la combinación de las distintas fuentes ahora disponibles, pero, al mismo tiempo, se crea además una nueva necesidad de **mejorar la coordinación e integración** entre todas las partes para poder así habilitar una producción y reutilización de los datos más eficiente y obtener así las mejoras esperadas en el ámbito social, económico y científico. Por tanto, el

reto actual consiste en la búsqueda de **nuevos modelos y relaciones de colaboración** entre las distintas partes que den una respuesta adecuada a los desafíos existentes a la hora de conseguir una **mejor gestión y mayor disponibilidad** de datos, beneficiando finalmente a todas las partes y facilitando su reutilización en mayor medida.

Para tener una visión más completa de la situación y conocer más en detalle quiénes son estos nuevos sectores productores de datos, cuáles son sus motivaciones y qué nuevos modelos de colaboración están surgiendo entre ellos, se ha elaborado este informe, que se divide en dos partes:



En la primera parte del informe haremos un recorrido por los distintos **sectores productores de datos** – dejando el sector público aparte por un momento – y las fuentes de datos alternativas que nos ofrecen, mostrando sus **motivaciones**, sus particularidades y varios ejemplos.



En la segunda parte del informe nos centraremos en analizar cuáles son los **modelos de colaboración** que pueden surgir entre el sector público y otros sectores para dar lugar a actuaciones conjuntas que consigan mejorar la gobernanza, disponibilidad y reutilización final de los datos, mostrando nuevamente las diferencias y semejanzas entre cada uno de ellos y distintos ejemplos de cómo se han puesto en práctica hasta el momento.

PARTE I: OTROS SECTORES PRODUCTORES DE DATOS

El discurso de los datos abiertos ha estado frecuentemente enfocado casi en exclusividad hacia los datos gubernamentales, un sector en el que la apertura de datos es fácilmente justificable por cuestiones de mejora en la transparencia, la eficiencia y la rendición de cuentas. Sin embargo, muchas veces se han obviado otras posibles motivaciones que podrían convertir a otros sectores en valiosas fuentes de datos alternativas disponibles a la vez para su reutilización.

Así pues, el objetivo de esta primera sección del informe es hacer un **recorrido por esas otras posibles fuentes de datos provenientes de distintos sectores más allá del sector público** – incluyendo el sector privado, el sector académico y de la innovación, el conocido como *tercer sector*, y la propia ciudadanía – analizando las posibles motivaciones en cada caso e incluyendo ejemplos reales que nos sirvan para dejar totalmente claro su potencial a la hora de proporcionarnos también datos de alto valor.

La siguiente imagen recoge estos publicadores de datos, distintos al sector público:



1.1. DATOS DESDE EL SECTOR PRIVADO

El sector privado es probablemente el mayor generador y gestor de datos existente en la actualidad, aunque la gran mayoría de esos datos no están disponibles de forma pública. Las empresas cuentan en general con una serie de incentivos comunes a la hora de compartir sus datos, entre los que se encuentran principalmente:

Incentivos para que las organizaciones privadas abran sus datos



GENERACIÓN DE BENEFICIOS



RECIPROCIDAD



INVESTIGACIÓN E INNOVACIÓN



MEJORA DE LA REPUTACIÓN

Datos abiertos más allá del sector público, datos.gob.es

- **Generación de beneficios:** la apertura, cesión o intercambio de los datos corporativos no tiene por qué ser siempre de carácter gratuito. En algunos casos esos datos podrían proporcionarse a cambio de una compensación económica, o incluso a modo de nuevo servicio, contribuyendo así a generar ingresos adicionales para la empresa.
- **Reciprocidad:** las empresas pueden compartir sus datos con otras entidades en un intercambio para beneficio mutuo, buscando así obtener acceso a otras fuentes de datos que pueden ser importantes para sus propias decisiones de negocio.
- **Investigación e innovación:** la apertura de los datos corporativos puede generar nuevas respuestas a problemas particulares, proporcionando a las empresas nuevas perspectivas más allá de sus propias fronteras que no podrían haber obtenido de otra manera, y obteniendo acceso a otras habilidades que tal vez no estuviesen disponibles en la propia empresa hasta ese momento.

- **Mejora de la reputación:** la cesión de datos por parte de las empresas, especialmente cuando el objetivo que se persigue es un bien común para la sociedad, puede contribuir también a la mejora de su reputación y de su imagen corporativa, lo que a su vez es una nueva puerta que se abre a la hora de atraer a potenciales nuevos usuarios, clientes, empleados e inversores.

En los siguientes apartados haremos un recorrido por algunas **motivaciones más específicas para la apertura de datos en el sector privado**, así como por varios ejemplos para cada una de ellas. La siguiente figura resume estas motivaciones, que se verán detalladamente a continuación.



1.1.1 – Regulación en sectores específicos

Existen ciertas regulaciones que afectan a distintas áreas específicas del sector privado y que, de un modo u otro, crean nuevas **obligaciones de publicación de**

datos y amplían nuestras posibilidades de acceder a fuentes de datos alternativas y complementarias a las del sector público.

Así por ejemplo, en el sector bancario las entidades que operan en territorio Europeo, y particularmente aquellas que están especialmente designadas por su relevancia, tienen la [obligación de reportar periódicamente a la entidad supervisora del Banco Central Europeo](#) una serie de datos entre los que se incluyen estadísticas generales y datos más específicos sobre balances, beneficios, capital, calidad de los activos, financiación o liquidez. Unos datos que la entidad supervisora pone también [a disposición del público](#), aunque únicamente de forma agregada.

Otro ejemplo que afecta tanto al ámbito público como al privado en el sector comercial es el de la [regulación Europea sobre comercio entre estados miembro](#), que obliga también a proporcionar una serie de datos sobre dichas transacciones, incluyendo información detallada sobre el tipo, valor y cantidad de las mercancías afectadas, los países de origen y llegada o la modalidad de transporte, entre otros. Estos datos estarán luego también [disponibles a través de Eurostat](#).

1.1.2 – Transparencia institucional

Desde la entrada en vigor de la [Ley de Transparencia, Acceso a la Información Pública y Buen Gobierno](#) todas las entidades de la administración general del estado, las administraciones de las comunidades autónomas y la administración local, así como otras entidades de carácter público, están obligados por ley a publicar una serie detallada de información de forma periódica, actualizada, clara y gratuita.

Sin embargo, algo que suele pasar más desapercibido es que dicha ley obliga también a **cualquier entidad privada que perciba ayudas o subvenciones públicas** en determinadas cuantías a publicar una serie de información institucional, económica, presupuestaria y estadística de forma proactiva, entre la que se incluye:

- Información relativa a sus funciones y la normativa aplicable.
- Estructura organizativa.
- Retribuciones de las personas con mayor responsabilidad.

- Contratos y convenios, incluyendo objeto, duración, importe y procedimientos.
- Subvenciones y ayudas públicas.
- Estadísticas del volumen presupuestario, incluyendo las principales partidas.
- Cuentas anuales e informes de auditoría y fiscalización.

Además, más allá de estos requisitos mínimos, la **transparencia económica y fiscal** es también un hábito que beneficia a las empresas a la hora de **generar confianza y atraer una mayor inversión**. Por tanto, no es de extrañar que según el último [informe de contribución y transparencia fiscal](#) las prácticas de transparencia y responsabilidad fiscal de las empresas del IBEX 35 no haya dejado de mejorar en los últimos años, aumentando así la cantidad y el detalle de la información interna que las empresas acaban haciendo pública.

1.1.3. Responsabilidad social corporativa

La responsabilidad social corporativa (RSC) hace referencia a un conjunto de **prácticas, estrategias y sistemas de gestión empresarial cuyo objetivo es conseguir un mejor equilibrio entre las dimensiones económica y social de las empresas**, buscando el balance adecuado y justo entre los intereses de la propia empresa y los de la sociedad en su conjunto.

En general las acciones de RSC son de carácter voluntario, aunque es frecuente que sigan las buenas prácticas y recomendaciones establecidas por organizaciones de referencia como la [Global Reporting Initiative](#) o los [Objetivos de Desarrollo Sostenible](#). Sin embargo, la reciente entrada en vigor de la conocida como [ley de información no financiera](#) abre todo un nuevo mundo de posibilidades en esta área al establecer que las grandes compañías – a partir de ciertos márgenes de empleados y facturación – deben compartir cierta información mínima respecto al impacto que tienen sus actividades en una serie de ámbitos como son:



Cuestiones económicas no financieras: incluyendo información sobre el modelo de negocio, relaciones comerciales, riesgos

vinculados a la actividad, prevención de corrupción y otros indicadores no financieros.



Cuestiones sobre desarrollo: incluyendo compromisos relacionados con el desarrollo sostenible y aplicación de políticas sociales y ambientales a las relaciones con proveedores y consumidores.



Cuestiones medioambientales: incluyendo actuaciones relativas a la contaminación, gestión de residuos, uso sostenible de recursos, cambio climático y protección de la biodiversidad.



Cuestiones sociales y de personal: incluyendo medidas relativas a la organización del trabajo, salud, seguridad, relaciones sociales, formación, accesibilidad, igualdad y diversidad.



Cuestiones relacionadas con los derechos humanos: incluyendo cumplimiento de los convenios internacionales y medidas de prevención y gestión de abusos y discriminación.

Según [el último informe del Observatorio de Responsabilidad Social Corporativa](#), por el momento la ley de información no financiera no está dando lugar a una mejora substancial en la información que proporcionan las empresas del IBEX 35. Sin embargo, es de esperar que esta situación mejore de forma sensible próximamente según nos vayamos acercando al límite del plazo para la implantación total de dicha ley en el año 2021.

Por otra parte, no hay duda de que existe también una creciente conciencia global en material social, económica y medioambiental que está forzando a las organizaciones corporativas a mostrar claramente su responsabilidad ya no sólo hacia sus accionistas, sino también hacia empleados, clientes y las comunidades en las que se implantan y gracias a las cuales prosperan. Esto da lugar por ejemplo a que la gran mayoría de las empresas del S&P 500 estén también publicando sus propios informes de responsabilidad social corporativa (en ausencia de obligación legal similar), e

incluso yendo en ocasiones mucho más allá con iniciativas temáticas más completas como el [portal de movilidad de Uber](#) o [la iniciativa medioambiental del operador de ferrocarriles Francés SNFC](#).

1.1.4– Filantropía de los datos

La filantropía de los datos (corriente conocida también como *data charities* o *data for good*) es una nueva forma de **altruismo** mediante la cual **empresas privadas deciden ceder una parte de sus datos a terceros de forma temporal o permanente persiguiendo contribuir a un bien social común** que en principio no tiene ninguna relación con su actividad principal, todo ello sin tener ninguna obligación particular al respecto. A cambio, dichas empresas pueden obtener un valor añadido no sólo por la imagen positiva que se obtiene, sino porque esa mejora social normalmente conlleva también una mejora general del entorno dentro del cual funciona su negocio que les beneficiará directamente.

Entre los beneficios de este nuevo tipo de filantropía se encuentran también el de ayudar a resolver problemas que no pueden ser abordados exclusivamente desde los datos públicos existentes, además de contribuir a la toma de decisiones informada y a la construcción de políticas públicas basadas en la evidencia e inspirar también nuevas soluciones innovadoras a la hora de dar respuesta a distintos problemas sociales.

Por ejemplo, Facebook cuenta con un programa “[Data for Good](#)” a través del cual aprovechan el potencial de los datos proporcionados por sus miles de millones de usuarios de base para – después del correspondiente proceso de *anonimización* – explorar su potencial para ayudar a las comunidades menos favorecidas, grupos de riesgo, minorías o personas en situaciones vulnerables en general. Gracias a ello pueden ofrecernos herramientas de datos muy útiles como [mapas de zonas catastróficas](#), de [redes de distribución eléctrica](#), de [densidad de población](#) o para la [prevención de enfermedades](#).

Otro ejemplo similar es el de [la línea de actividad de “Data for Good” de LUCA](#) – la unidad de Telefónica especializada en Big Data e Inteligencia Artificial – a través de la cual se hace uso de los datos para proyectos sociales, humanitarios y ambientales en línea con los objetivos de desarrollo de la Organización de las Naciones Unidas (SDGs).

1.1.5 – Servicios comerciales de intermediación de datos

Un **mercado de datos** (frecuentemente conocidos como *data markets* o *data marketplaces*) es un **servicio comercial de intermediación que ofrece acceso previo acuerdo a varios tipos de datos procedentes de diversas fuentes** (incluyendo datos públicos y privados) y que, generalmente, ofrece también algunas **herramientas adicionales** o incluso la posibilidad de *moldear* los datos según las necesidades específicas de sus clientes a través de servicios de valor añadido como la unificación de formatos, interoperabilidad, indicadores de calidad, analítica o acceso a través de interfaces avanzados. Ejemplos de este tipo de mercados de datos son [IOTA](#), [Qlik Data Market](#), [Quadrant](#), [Refinitiv](#), [Adveneo](#) o [Dawex](#).

Algunos de los tipos de datos ofertados más comunes incluyen inteligencia de negocios, publicidad, demografía, información personal, investigación y datos de mercado. Existen también mercados de datos más especializados en algunas categorías de datos en particular, por ejemplo, datos geográficos o [datos humanitarios](#). Los datos que se venden en estos mercados son utilizados por empresas de todo tipo, desde gobiernos a agencias de inteligencia de mercado, pasando por distintos tipos de analistas.

Según Edd Dumbill, analista de O'Reilly Radar, [la historia de los mercados de datos se remonta a mediados del siglo XIX](#), cuando Paul Reuter, fundador de los medios de comunicación de Reuters, comenzó a publicar los precios de intercambio bursátil entre las bolsas de París y Londres. Sin embargo, ha sido en los últimos diez años cuando este tipo de mercados han proliferado al albor del crecimiento de las nuevas tendencias focalizadas en los datos, como el *big data* o el [Internet de las cosas](#), ya que la cantidad de datos disponibles ha aumentado exponencialmente y el apetito de

las empresas por esos datos ha crecido también en la misma medida, por lo que en la actualidad los datos se han constituido como un activo de gran valor y demanda. Por ejemplo, se estima que para el año 2030 más de 1 millón de organizaciones estarán monetizando sus datos IoT y se llevarán a cabo transacciones de más de 12 exabytes de datos cada día.

En estos mercados los proveedores o propietarios originales de los datos autorizan al intermediario para que licencie su información en base a unos términos y condiciones definidas. Por tanto, es importante no confundirlos con los también populares data brokers, como Quantcast o Acxiom, cuya función consiste en recopilar el mayor número de datos posible, generalmente de carácter personal (que nunca deberían estar presentes en un mercado de datos al uso) o relacionados con el consumo, con el objetivo de caracterizar ciertos grupos o individuos con los que nunca han mantenido ningún tipo de relación comercial o de servicio, y haciendo uso en ocasiones de métodos bastante opacos y éticamente (o incluso legalmente) cuestionables.

1.2. - DATOS DEL SECTOR ACADÉMICO Y DE LA INNOVACIÓN

El sector académico ha sido históricamente otro gran productor de datos debido a su continua actividad científica. Gracias a los avances tecnológicos que han llegado también a la comunidad científica, toda esa vasta producción de datos es hoy en día mucho más accesible al público de lo que lo era con anterioridad. A esto hay que añadir que los datos se han convertido en una herramienta casi imprescindible para la innovación, por lo que cualquier actividad u organización con un alto componente innovador se convierte automáticamente en otra potencial gran fuente de datos.

A través de los siguientes apartados veremos distintas motivaciones - recogidas en la siguiente imagen- a la hora de facilitar el acceso a los datos provenientes de estos sectores, así como nuevamente varios ejemplos reales y actuales para cada una de ellas.

¿Por qué motivos comparte el sector académico y de la innovación datos?



ACCESO ABIERTO A LOS RESULTADOS CIENTÍFICOS

Los resultados de investigaciones son accesibles y cualquiera puede replicar/acceder a un estudio.



DATOS DE DOMINIO PÚBLICO

Los datos son un recurso compartido que se explota por el beneficio común.



LABORATORIOS DE EXPERIMENTACIÓN CON DATOS

Crear soluciones innovadoras y prácticas ante cualquier problema.



SISTEMAS COLABORATIVOS DE DATOS

Data Hubs para trabajar fluidamente de forma colaborativa

Datos abiertos más allá del sector público, datos.gob.es

1.2.1 – Acceso abierto a los resultados científicos

Existe una cantidad enorme de investigaciones científicas realizadas en universidades e instituciones de todo el mundo – que además hay que recordar que en muchas ocasiones son también entidades públicas y/o costeadas con recursos públicos. Sin embargo, el acceso a este conocimiento que se produce como un bien público suele estar restringido, y sólo se concede a aquellos a los que se les concede permiso a través de su afiliación universitaria, o mediante la adquisición de derechos de acceso a través de revistas académicas o artículos individuales.

El **Open Access**, o publicación científica en acceso abierto, nace como solución a este problema y para conseguir que **los resultados de la investigación sean más accesibles a todos** los agentes sociales, contribuyendo a una **ciencia mejor y más eficaz y a mejorar así la innovación** en los sectores público y privado. Los orígenes de este movimiento se remontan al compromiso social y científico alcanzado a nivel internacional a través de las siguientes declaraciones:

- [Budapest Open Access Initiative](#) (2002): Donde se aboga por el libre acceso a la literatura científica a través de Internet.

- [Declaración de Bethesda sobre la publicación de acceso abierto](#) (2003): Donde se recomienda el archivo de los trabajos científicos en repositorios de acceso abierto.
- [Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities](#) (2003): Donde se ratifican las dos declaraciones anteriores, se apoya abiertamente el movimiento *Open Access* y se compromete a las instituciones a favorecer e impulsar esta vía.

Según dichas declaraciones, el acceso abierto se define con las siguientes características: **digital, online, gratuito y libre de la mayoría de las restricciones de derechos de autor y licencias**. Esto facilita que cualquiera pueda leer, descargar, copiar, distribuir, imprimir, buscar o enlazar los textos de los artículos científicos y los datos que los acompañan para poder usarlos con cualquier otro propósito legítimo.

El movimiento pro acceso abierto ha conseguido rápidamente un gran respaldo internacional y gracias a ello hoy en día existen grandes repositorios (y *meta-repositorios*) de contenido científico (incluyendo datos) de acceso abierto, tales como [SciELO](#) (Scientific Electronic Library Online), [ROAR](#) (Registry of Open Access Repositories), [DOAJ](#) (Directory of Open Access Journals) u [OAPEN](#) (Open Access Publishing in European Networks) entre otros.

En España el [Real Decreto 99/2011](#), por el que se regulan las enseñanzas oficiales de doctorado, establece la obligatoriedad por parte de la Universidad correspondiente de archivar una copia de tesis doctorales en formato electrónico abierto en un repositorio institucional, mientras que la [Ley 14/2011 de la Ciencia, la Tecnología y la Investigación](#) establece también la obligatoriedad de consignar en repositorios de acceso abierto dentro de su propio campo de conocimiento una versión digital de la versión final de aquellos proyectos que hayan sido financiados (mayoritariamente) con fondos de los Presupuestos Generales del Estado. [RECOLECTA](#) es la plataforma que agrupa a todos los repositorios de ciencia abierta nacionales, con más de dos millones de documentos indexados hasta la fecha.

La Comisión Europea por su parte ha estado también llevando a cabo [un proyecto piloto \(ORD\) sobre el acceso abierto](#) a los datos de investigación en el programa Horizonte 2020 como parte de su objetivo a largo plazo de conseguir que los datos de investigación sean siempre *abiertos por defecto*. Gracias a este piloto, todos los proyectos que reciban financiación del programa deberán asegurarse de que los artículos y datos que publiquen estén disponibles de forma abierta y gratuita.

1.2.2 – Datos de dominio público

A lo largo de la historia las comunidades han venido compartiendo y gestionando de forma común ciertos recursos de especial relevancia por el propio **beneficio de la comunidad**, como por ejemplo los bosques, la caza o las zonas de pastoreo, siempre de una forma *autorregulada* y dentro de lo que se conoce como [dominio público](#).

El enfoque del uso de los datos basado en el dominio público (o *data commons*) sigue una aproximación similar a través de la cual **los datos se erigen como un nuevo recurso compartido que se explotarán por el beneficio común**. Una vez que se ha establecido una comunidad de interés en torno a una serie de datos hay dos actuaciones principales que debemos llevar a cabo para llegar a crear un *data commons* de forma efectiva:

1. El diseño colaborativo de las reglas y protocolos de uso

Elaborando unos principios y protocolos de actuación (tanto tecnológicos como sociales) también conocido como "*contrato de comunidad*" que establezcan el acuerdo común y rijan la reutilización de los mismos. La manera en la que se desarrollen estos protocolos es tan importante como los propios protocolos, siendo la participación entre iguales clave para el éxito.

2. Puesta en marcha del *Data Commons*

Implementando soluciones específicas de reutilización de datos que utilicen los protocolos diseñados como base para las relaciones con la comunidad de usuarios y consigan proporcionar un alto valor para ellos a través de sus propios datos.

Existen además una serie de principios básicos que serán fundamentales a la hora de construir un modelo de integración y reutilización de datos basado en el dominio público:

- Los datos deberán ser un **recurso de uso común y para el bien común**.
- El valor añadido vendrá dado por la **reutilización o la integración** con otros datos.
- Debe garantizarse el **uso correcto** de los datos, de acuerdo a las reglas establecidas.
- El sistema debe estar diseñado para la **participación inclusiva** de todos sus miembros.
- El **control** final debe estar en las manos de **aquellos más afectados** por las decisiones que se vayan a tomar en torno a los datos.
- Las **reglas de uso y participación** deben diseñarse para facilitar la escalabilidad.

Un buen ejemplo de la aplicación del dominio público a los datos es el del proyecto [Genomic Data Commons](#), cuyo objetivo es el de proporcionar a la comunidad científica que investiga sobre el cáncer un repositorio unificado de datos sobre estudios genómicos que sirva para elaborar tratamientos de mayor precisión.

1.2.3 – Laboratorios de experimentación con datos

Los conocidos como laboratorios de datos (o *Data Labs*) no son más que entornos de **experimentación que generalmente buscan dar soluciones innovadoras**, pero de carácter práctico, a cualquier problema social apremiante a través del análisis en profundidad de los datos y con una aproximación que va desde el problema a la solución (*bottom-up*). En estos laboratorios los datos son los protagonistas a la hora de establecer enfoques a medida basados en el análisis de las realidades más próximas al problema en cuestión. Para ello, suelen trabajar de forma muy cercana

con las distintas comunidades a las que afectan esos problemas, estableciendo conexiones de larga duración y recopilando datos de nueva creación para rellenar los vacíos existentes.

Este tipo de entornos experimentales de datos son también bastante populares en el entorno académico, contando con ejemplos como el del [Living Lab del MIT](#) que trabaja en áreas tan diversas como la salud y el bienestar, el análisis de los patrones sociales y el comportamiento humano o el transporte y la movilidad; el [Stanford Data Lab](#) en áreas como gobierno, sanidad o pobreza; o también el [Pew Research Center](#), que trabaja en [proyectos relacionados con las ciencias sociales](#) en general. Otro ejemplo es el del [Global Change Data Lab](#) y su proyecto de referencia [Our World in Data](#), que está íntimamente relacionado con la universidad de Oxford y se centra en el análisis de los cambios producidos por un mundo cada vez más global.

No obstante, existen también otros ejemplos fuera del ámbito puramente académico, como [el Data Lab promovido por el programa de centros de innovación del gobierno de Escocia](#) o el [SDG Data Lab](#) promovido por la *iTech Mission* para la monitorización del avance en los objetivos mundiales de desarrollo sostenible.

1.2.4 – Sistemas colaborativos de datos

Estos sistemas centralizados colaborativos o almacenes de datos (*Data Hubs*) están pensados para compartir los datos fácilmente independientemente de su origen, y poder así **trabajar fluidamente de forma colaborativa**. En ocasiones estos *hubs* nos ofrecerán además una serie de **herramientas básicas** para facilitarnos aún más nuestro trabajo con los datos. Sin embargo, no debemos confundir el producto técnico (la plataforma donde se alojan los datos y/o herramientas que la acompañan) con el ecosistema de datos y la comunidad que se crea a su alrededor y que a la par constituye el verdadero valor de estos *Hubs*.

Así podemos encontrar cientos de *Data Hubs* con grandes cantidades de datos disponibles ya para su uso, desde algunos que abarcan conocimientos generales

como la [DBPedia](#) o [Lexvo](#), hasta otros más específicos centrados en áreas concretas de conocimiento, como:



Geografía: [Geonames](#), [Ocean Drilling](#), o [Pleiades](#).



Biología: [Drugbank](#), [Bioportal](#) o [Eunis](#).



Lingüística: [PanLex](#), [Phoible](#) o [BabelNet](#).



Multimedia: [DBTune](#), [Chronicling America](#) o [Between our Worlds](#).



Bibliografía: [The European Library](#), [Proyecto Gutenberg](#) o [The British National Bibliography](#).

1.3 – LOS DATOS DEL TERCER SECTOR

El **tercer sector** hace referencia simplemente a aquel sector de la economía que no está directamente comprendido ni en el sector privado ni en el público. Las entidades de este sector – generalmente sin ánimo de lucro – surgen en su mayoría a través del empuje de la ciudadanía o por el interés de algunas empresas en involucrarse con acciones más concretas en algunos ámbitos sociales en los que están interesados.

Si bien se podría considerar que el volumen de datos generado en este sector es significativamente menor que el de otros sectores que hemos visto, hay que tener también en cuenta que el tercer sector se ha convertido en un actor muy relevante en el desarrollo social y que además, en muchas ocasiones, cubre necesidades únicas a las que nadie más está dando una respuesta adecuada. Por ese mismo motivo, e independientemente del volumen de datos del que estemos hablando, los datos provenientes de este sector generalmente cuentan con un alto valor por su carácter único.

Al igual que en los casos anteriores, en los siguientes apartados veremos distintas motivaciones a la hora de proporcionar acceso a los datos de este sector acompañadas de los correspondientes ejemplos, como muestra la siguiente figura:

¿Por qué motivos comparte el tercer sector datos?



TRANSPARENCIA EN LA GESTIÓN

Mejorar la coordinación, la rendición de cuentas y la eficacia.



DATOS PARA FOMENTAR EL DESARROLLO

Definir intervenciones y evaluar su progreso/efectividad.



DATOS PARA LA COLABORACIÓN MULTILATERAL

Unir fuerzas para atajar las problemáticas de los países.

Datos abiertos más allá del sector público, datos.gob.es

1.3.1 - Transparencia en la gestión

En la economía social y solidaria una mayor transparencia no es tan sólo una buena práctica, sino que, debido a los especialmente escasos recursos con los que usualmente cuentan estas organizaciones es también un paso esencial para **mejorar la coordinación, la rendición de cuentas y la eficacia**. Así pues:



Los **gobiernos de los países** en desarrollo necesitan información actualizada sobre las organizaciones humanitarias y de desarrollo que operan en su país, para poder trabajar con ellas de manera más eficaz.



Los **donantes y las agencias multilaterales** necesitan entender cómo y dónde sus inversiones podrían tener el mayor impacto.



Las **organizaciones de la sociedad civil** que prestan servicios necesitan ver qué está sucediendo en cada lugar y momento, para mejorar la coordinación y evitar la duplicación de esfuerzos.



La **ciudadanía** y los **periodistas** necesitan también esa información para que sus gobiernos rindan cuentas sobre sus gastos y el buen uso de los recursos.

No es por tanto de extrañar que algunas de las iniciativas pioneras y más completas en cuanto a transparencia en el funcionamiento de las organizaciones hayan surgido precisamente en el tercer sector, abriéndonos así las puertas a todo un valioso conjunto de datos sobre su funcionamiento interno.

Así por ejemplo, son más de 1.000 organizaciones del tercer sector las que han estado utilizando el [estándar](#) de la [International Aid Transparency Initiative](#) (IATI) durante los últimos 10 años para compartir detalles sobre sus presupuestos, áreas de actuación y resultados. Gracias al amplio conjunto de herramientas disponibles para [facilitar su uso](#), estos datos se han convertido ya en una valiosa herramienta a la hora de planificar la ayuda al desarrollo.

Otro ejemplo similar es el del [estándar 360Giving](#), cuya misión es apoyar a los donantes del Reino Unido para que publiquen información sobre a quién, dónde y qué financian en un formato abierto y estandarizado con el objetivo de construir una imagen completa y fidedigna del panorama de la financiación y poder así aumentar su impacto.

1.3.2 - Datos para fomentar el desarrollo

Los datos se han venido utilizado ampliamente en el desarrollo internacional desde hace muchos años. Gracias a ellos las organizaciones de desarrollo han podido **definir las intervenciones** más apropiadas y relevantes, **evaluar el progreso** de los proyectos a lo largo del tiempo **y evaluar la efectividad** de la agenda global de desarrollo. La digitalización ofrece ahora múltiples nuevas e interesantes oportunidades para el desarrollo al permitir usar esos mismos datos de manera más eficiente para fortalecer los procesos de toma de decisiones y mejorar la prestación de servicios, conseguir una participación ciudadana más significativa y aumentar la capacidad de respuesta en los servicios humanitarios. Los datos se han convertido en

definitiva en uno de los ingredientes clave que pueden contribuir a resolver los retos del desarrollo global, junto a las organizaciones, comunidades y líderes adecuados.

Unos de los principales actores involucrados en la producción, gestión y análisis de los datos para el desarrollo han sido históricamente los Bancos Internacionales de Desarrollo. Así, no es de extrañar que tanto el [Banco Iberoamericano de Desarrollo](#) (BID), como el [Banco Asiático de Desarrollo](#) y el [Banco Africano de Desarrollo](#) cuenten en la actualidad con sus propias plataformas de datos como herramientas clave en su lucha por reducir la pobreza y la desigualdad en sus respectivas zonas de actuación. Por supuesto, otras grandes organizaciones internacionales cuentan también con sus propios bancos de datos específicamente enfocados al desarrollo, como es el caso del [Banco Mundial](#), [USAid](#) o la [OECD](#). Existen además otras iniciativas interesantes fuera de las grandes organizaciones de referencia, como por ejemplo el [Development Data Hub](#) mantenido por Development Initiatives o [los proyectos de datos para el desarrollo llevados a cabo por el Development Data Lab](#).

Por último, no nos podemos olvidar de la que probablemente sea la base de datos de referencia para el mundo del desarrollo en la actualidad. Se trata de la plataforma de indicadores para la evaluación de los [objetivos de desarrollo sostenibles](#) (SDGs) promovidos por las Naciones Unidas, ya que contiene más de un millón de observaciones sobre el avance internacional hacia la consecución de dichos objetivos, que son precisamente los marcarán la evolución de la agenda de desarrollo global durante la próxima década.

1.3.3 – Datos para la colaboración multilateral

Los organismos multilaterales son organizaciones en las que generalmente participan múltiples naciones que trabajan de forma colaborativa con la misión de unir fuerzas para atajar las problemáticas de los países que las integran. Por tanto, este tipo de organizaciones resultan fuentes de datos muy valiosas (o incluso las más valiosas) a la hora de tener una perspectiva global sobre ciertos temas concretos. Algunas de las organizaciones que pueden resultar más útiles en este aspecto son:



BANCO MUNDIAL
BIRF • AIF

El **Banco Mundial** ha construido a lo largo de los años lo que podría considerarse como [uno de los catálogos de datos globales más completos y funcionales que pueden encontrarse en la actualidad](#) y que se puede explorar a través de [distintas variables](#), incluyendo [colecciones temáticas](#) y [bases de datos específicas](#). Además ha desarrollado otras herramientas ampliamente utilizadas, como su popular [catálogo de indicadores](#), el [banco de datos](#) o el [archivo de microdatos](#).

Por otro lado, cuenta también con el más completo [juego de herramientas](#) para que los desarrolladores puedan sacar mejor partido de los datos, así como su respectivo catálogo de documentación.



La **Organización de las Naciones Unidas** (ONU) nos proporciona acceso a 32 bases de datos con un total de más de 60 millones de registros a través de su [catálogo online](#), e incluyendo también datos de muchas de sus organizaciones asociadas como la Unión Internacional de Telecomunicaciones (ITU), la Organización Mundial Meteorológica (WMO), la Organización Laboral Internacional (ILO) o la Organización Mundial de Turismo (UNWTO). A todos estos datos se podrá acceder también a través de una [interfaz de programación](#) (API) completamente documentada y que utiliza el estándar SDMX para el modelado de los datos.

Cuenta además con las [bases de datos](#) de su división estadística y múltiples [indicadores](#) con los datos sobre la gestión interna del propio organismo.



Organización Mundial de la Salud

La **Organización Mundial de la Salud** (OMS) nos ofrece su [Observatorio Mundial de la Salud](#) como puerta de acceso a las principales estadísticas mundiales e indicadores relacionados con la salud. Así, podremos recorrer sus bases de datos a través de [distintas temáticas](#) relacionadas con la salud, buscando [indicadores](#)

específicos o centrándonos en un país en concreto. Cuentan también con un interfaz de acceso automatizado (API) para facilitarnos la reutilización.

Por otro lado, tendremos igualmente acceso a múltiples informes a través de los cuales se desgranar los datos y se analizan las conclusiones extraídas a partir de ellos.



La **Organización para la Cooperación Económica y el Desarrollo** (OECD) cuenta en su catálogo online con más de 800 bases de datos que cubren temáticas tan variadas como *agricultura, desarrollo, economía, educación, energía, medio ambiente, finanzas, sanidad o tecnología*. Además, una cuarta parte de ellas ofrece también acceso mediante interfaces de programación (APIs) totalmente documentadas que facilitan su consulta automatizada por parte de programadores.

Cabe destacar por otro lado su completo buscador de estadísticas, que nos ofrece también una completa documentación sobre cómo se elaboran. Además de otros complementos útiles como visualizaciones o herramientas de comparación.



El **Fondo Monetario Internacional** (FMI) publica una serie completa de datos económicos globales, así como datos detallados sobre tipos de cambio y otros indicadores económicos y financieros. También proporciona manuales, guías y otros materiales que documentan sus propias prácticas estadísticas y las de los países miembros.

Nuevamente, los datos vienen también acompañados por una completa herramienta de visualización que nos permite hacer comparaciones fácilmente.

1.4 - LOS DATOS DE LAS PERSONAS

Para terminar esta sección no podemos dejar de mencionar un área que en cierta medida se está revelando como la fuente de datos más codiciada en la actualidad, aunque no siempre con consecuencias positivas: Los datos personales.

Y es que el modelo mayormente predominante a la hora de explotar los datos personales hoy en día se centra en tres pilares fundamentales: (1) la propiedad de los datos; (2) el consentimiento o control mínimo por parte de las personas; y (3) el beneficio comercial a corto plazo. Este modelo nos ha llevado a una industria creciente y rentable en torno a la recopilación, integración y monetización de los datos, pero al mismo tiempo ha dado lugar a efectos laterales no deseados, principalmente en cuanto a todo aquello que afecta a la privacidad.

Así pues, en este apartado veremos algunas de las motivaciones alternativas existentes a la hora de trabajar con los datos que proceden directamente de la ciudadanía de una forma más equitativa a través de dos modelos principales: la autogestión y la co-creación.

¿Por qué motivos los ciudadanos comparten datos?



AUTOGESTIÓN DE DATOS PERSONALES

Espacio personal de almacenamiento que permite controlar quién y para qué accede a los datos.



REPOSITORIOS DE DATOS COLABORATIVOS

Datos generados por las personas de forma colaborativa.

Datos abiertos más allá del sector público, datos.gob.es

1.4.1 – Autogestión de los datos personales

Los sistemas de gestión de datos personales (*personal data stores* o *PDSs* en inglés) surgen como alternativa a los modelos de gestión centralizados dominantes en la

actualidad que concentran nuestros datos en sistemas independientes y aislados. En su lugar, los sistemas de gestión de datos personales se caracterizan por **devolver a las personas la capacidad de decidir sobre sus datos**, ofreciéndoles un **espacio personal de almacenamiento** (que puede perfectamente seguir siendo distribuido) y una **gestión unificada e interoperable** de los datos que les permite **controlar** de forma directa y en todo momento **quién** tiene acceso a los mismos **y para qué**, pudiendo para ello establecer permisos específicos personalizados.

Este modelo de gestión no sólo ofrece ventajas desde el punto de vista del fortalecimiento de la privacidad y los derechos digitales, sino que también abre todo un nuevo espectro de oportunidades para la innovación en servicios, gracias a que facilita la combinación (consentida) de un gran número de datos que hasta el momento estaban aislados en plataformas específicas.

Uno de los ejemplos más recientes de estos sistemas es [SoLiD](#), una propuesta dirigida por el creador de la Web – Tim Berners-Lee – para dar respuesta al bloqueo de los datos personales debido a la falta de interoperabilidad entre aplicaciones existente actualmente. El objetivo de SoLiD es crear todo un ecosistema para llevar a cabo una nueva descentralización mediante la cual las personas puedan elegir dónde se almacenan sus datos y quién puede acceder a ellos, fomentando también la innovación gracias a evitar el bloqueo que supone la falta de interoperabilidad entre las plataformas de datos actuales. Otros ejemplos de plataformas similares son [Digi.me](#), [openPDS](#), o la española [Privacy Cloud](#).

Sin embargo, aunque el concepto de los sistemas de gestión de datos personales ha sido ya llevado a la práctica en repetidas ocasiones y existe toda una variedad de plataformas tanto gratuitas como comerciales desde hace algunos años, estos sistemas no han conseguido establecerse como una alternativa clara todavía, incluso en un entorno como el actual de creciente interés y preocupación por la privacidad de las personas.

Probablemente la principal barrera para la difusión de estos sistemas viene precisamente de esa adopción limitada hoy en día, lo cual contribuye a crear un círculo

vicioso en el que es muy difícil que se ofrezcan servicios interesantes de valor añadido que puedan atraer a nuevos usuarios precisamente hasta que no se consiga contar con una masa crítica mínima de personas. No obstante, el nuevo [derecho a la portabilidad de los datos personales](#) establecido por el Reglamento General de Protección de Datos Europeo (GDPR) puede contribuir significativamente a paliar este problema en un futuro cercano. Por ejemplo, gracias a ello ya están surgiendo iniciativas como el [Data Transfer Project](#), promovido por las grandes plataformas de contenidos de la industria TIC.

1.4.2 – Repositorios de datos colaborativos

Las fuentes de datos colaborativas surgen del concepto de ver a los **ciudadanos** no únicamente como productores o consumidores pasivos de información, sino también **como editores, usuarios e intermediarios activos de los datos**.

La utilización de un enfoque de **datos generados por las personas de forma colaborativa** puede ser eficaz en situaciones en las que no se dispone de datos de otras fuentes de utilidad para las partes interesadas afectadas, ya sea porque no son recolectados por el gobierno o el sector privado o simplemente porque, aunque pueda ser que realmente existan, éstas no son accesibles de cualquier modo. Asimismo, los datos generados activamente por las personas pueden ser también útiles para complementar y/o contrastar los datos ya existentes cuando éstos son incompletos, inexactos o no lo suficientemente oportunos.

Entre este tipo de fuentes de datos colaborativas se encuentran algunas de las más completas, populares y reutilizadas fuentes de datos con las que contamos hoy en día, como por ejemplo [Open Street Maps](#) o [Wikidata](#) (que engloba a todos los proyectos de Wikimedia, incluida la Wikipedia), pero también otras mucho menos conocidas aunque igualmente útiles como [Open Food Facts](#) (base de datos colaborativa de productos alimentarios), [Open Africa](#) (el mayor repositorio colaborativo de datos abiertos en el continente), o [el Viajero](#) (plataforma de contenidos turísticos que incluye recomendaciones personales).

PARTE II - MODELOS DE COLABORACIÓN ENTRE ENTIDADES PÚBLICAS Y PRIVADAS

En esta segunda parte del informe nos dedicaremos a presentar distintos modelos de colaboración que pueden surgir entre el sector público y los distintos agentes productores de datos vistos en la sección anterior a la hora de trabajar juntos para aumentar la disponibilidad de datos. Varios de estos modelos -representados en la siguiente figura- no son una invención nueva del mundo de los datos, sino que simplemente están adaptados de otros modelos ya existentes para la regulación de relaciones similares en otros ámbitos de aplicación. Algunos de ellos pueden parecer también muy similares entre sí, pero en esos casos son precisamente las sutiles diferencias existentes lo que mejor los define.

Al igual que en la sección anterior, veremos también varios ejemplos de cómo se han ido poniendo en práctica cada uno de estos modelos hasta el momento.



2.1 – ACUERDOS DE COMPARTICIÓN DE DATOS

Un acuerdo de compartición de datos (*Data Sharing Agreement* o DSA) podría definirse de forma general como un **acuerdo entre dos o más entidades legales para el intercambio de datos o información de cualquier tipo**. Este tipo de procedimiento para la puesta en común de datos es relativamente frecuente y puede tomar distintas formas dependiendo de cada caso concreto, desde [consorcios entre múltiples organizaciones públicas y privadas para mejorar la movilidad y sostenibilidad de los sistemas de transporte](#) hasta [acuerdos entre el Instituto Nacional de Salud \(NIH\) de Estados Unidos y la nación del pueblo Navajo](#) en materia de salud infantil.

Según las necesidades específicas de las partes implicadas, esta puesta en común de datos puede adoptar distintas formas que van desde un intercambio recíproco de datos a la compartición con otras organizaciones, ya sea de forma puntual o más continua en el tiempo. También son bastantes frecuentes los acuerdos de compartición entre distintas partes de una misma organización (por ejemplo entre distintas agencias de un mismo gobierno).

Estos tipos de acuerdos puede afectar a datos de naturaleza muy variada, incluyendo datos personales u otros datos protegidos por derechos de propiedad intelectual y similares. Es por ello que dentro de estos acuerdos habrá que contemplar también la legislación general aplicable a los distintos tipos de datos involucrados (como por ejemplo la normativa de protección de datos), sin perjuicio de cualesquiera otros términos y condiciones que se acuerden particularmente, tales como la duración del acuerdo, cuestiones de confidencialidad, medidas de seguridad o restricciones de uso.

Dada la dificultad de establecer un marco normativo que pueda ser efectivo ante la gran variedad y divergencia de todos los posibles casos de uso existentes para el intercambio de datos entre organizaciones, la Comisión Europea ha optado en su lugar por publicar una serie de [pautas detalladas para el intercambio de datos](#) en forma de buenas prácticas y recomendaciones a seguir. Otros gobiernos han optado por publicar guías más específicas para atender mejor a casos más particulares, como el

[código de conducta a la hora de compartir datos personales](#) publicado en el Reino Unido a través de su Information Commissioner's Office (ICO) y cuyo objetivo principal es facilitar la correcta implementación de la normativa de protección de datos.

2.2 – DATA TRUSTS

El [Open Data Institute](#) (ODI) define a los *data trust* (*fideicomiso* o *fondo fiduciario* de datos) como **estructuras organizativas pensadas para facilitar la gestión independiente de los datos** y que se inspiran en la forma que toman los fideicomisos legales pero que no llegan a tener la misma consideración legal exactamente. Los *data trust* toman *prestado* ese concepto de permitir que una organización independiente tome el control de las decisiones sobre algo y lo aplican al mundo de los datos.

Como administrador de los datos, un *data trust* puede decidir quién tendrá acceso a esos datos, bajo qué condiciones y en beneficio de quién. Mientras que normalmente la misma organización que recopila y conserva los datos suele ser también la encargada de administrarlos, en el caso de los *data trust* se permite que sean otras organizaciones externas las que puedan tomar decisiones independientes sobre cómo se van a usar y compartir esos datos, según considere más apropiado en cada caso y en concordancia con los propósitos inicialmente previstos.

Los *data trust* se mantienen independientes de las organizaciones que poseen los datos y de los posibles usuarios de los mismos. Para lograr esta independencia es posible que tanto los titulares originales de los datos como los usuarios finales puedan en algún momento verse apartados de las decisiones que se toman sobre los mismos. Al mismo tiempo, **el *data trust* asumirá la responsabilidad jurídica vinculante de garantizar que los datos se compartan y utilicen en beneficio de un grupo particular de personas y organizaciones según lo acordado.**

Existen varios beneficios potenciales asociados a la adopción de esta figura intermediaria para la gestión de los datos. Por ejemplo, como organismo independiente, un *data trust* puede ayudar a equilibrar los puntos de vista e incentivos

de los distintos agentes implicados, que en varias ocasiones podrán ser diferentes, si no directamente contrapuestos. También, al tratarse de organismos más especializados, los *data trust* pueden ayudar a conseguir todos los beneficios generalmente asociados a compartir los datos, pero a un coste significativamente menor gracias precisamente a esa especialización.

El ODI ha estado poniendo este modelo a prueba recientemente a través de [una serie de experiencias piloto](#) realizadas conjuntamente con varias organizaciones públicas y privadas en el Reino Unido, con objetivos tan dispares como la gestión de los aparcamientos reservados a vehículos eléctricos, la minimización de desperdicios procedentes de excedentes alimentarios o el tráfico de animales y plantas.

2.3 – DATA COLLABORATIVES

El término *Data Collaboratives* (colaboración de datos) se refiere a un nuevo modelo de colaboración que va más allá del modelo tradicional de asociación público-privada (*public-private partnerships* o ppp), en el que **los participantes de diferentes sectores** (incluyendo empresas privadas, instituciones de investigación y agencias gubernamentales) **intercambian datos con el objetivo de ayudar a resolver los problemas públicos**.

El acceso a diferentes fuentes de datos nos permite entender los problemas públicos desde distintos ángulos. Mediante la creación de este tipo de colaboraciones de datos y la apertura de datos de distintos agentes, se pueden encontrar nuevas soluciones para combatir los problemas públicos. Así, los *Data Collaboratives* pueden contribuir a generar valor público de muy diversas formas, como por ejemplo:



En la respuesta a situaciones excepcionales o de emergencia: facilitando la comprensión del problema y las capacidades para rastrear las condiciones sobre el terreno en cada momento para ayudar así a ejecutar intervenciones más eficientes.



El diseño y la ejecución de los servicios públicos: ya que el acceso a nuevos conjuntos de datos da lugar a una modelización más

precisa de la población que permitirá una adaptación también más precisa y selectiva basada en la evidencia.



La creación y transferencia de conocimiento: puesto que una mayor disponibilidad de datos permite reducir los vacíos de conocimiento existentes y garantizar que aquellos que sean responsables de la toma de decisiones cuenten siempre con toda la información necesaria al respecto.



La predicción y anticipación: habilitando nuevas capacidades que contribuirán a que las instituciones puedan ser más proactivas poniendo en marcha mecanismos destinados a mitigar potenciales problemas o incluso evitarlos antes de que ocurran.



La evaluación del impacto: pudiendo monitorizar el impacto real de las políticas gracias a la mayor disponibilidad de información y tomar así las acciones correctivas necesarias a tiempo.

El GovLab, promotor del concepto de *data collaboratives*, nos ofrece [un amplio catálogo con docenas de ejemplos reales de colaboraciones en activo](#) en diferentes partes del mundo y abarcando áreas tan diversas como agricultura, educación, medio ambiente, educación, transporte o infraestructuras entre otras.

2.4 – DATA FELLOWSHIPS

Los *data fellowships* consisten generalmente en **estancias programadas en las que miembros del gobierno comparten tiempo y espacio con otras organizaciones privadas**. El objetivo principal de estas colaboraciones es **intercambiar conocimientos** a la vez que se trabaja de forma conjunta en resolver algún problema específico de interés público. Algunas ventajas de los programas de *fellowship* son:



Mejoran tanto las **capacidades personales** como las de la **organización**, gracias al proceso continuo de formación, mentorización y aprendizaje.



Proporcionan una **experiencia práctica** trabajando sobre **casos reales**.



Facilitan el desarrollo de nuevos **especialistas** dentro de la administración que contribuirán a la **difusión interna** de los conocimientos adquiridos.



Contribuyen a ampliar las **redes de colaboración** más allá de los círculos habituales.



Ayudan a **comprender mejor el ecosistema de los datos** y los diferentes agentes que forman parte del mismo.

Un ejemplo de este modelo es [el programa de data fellowship llevado a cabo por el gobierno de Australia](#), a través del cual se seleccionan una serie de empleados públicos para trabajar con organizaciones privadas con el objetivo de desarrollar sus capacidades y dar solución a problemas concretos en sus respectivas agencias de origen.

Por otra parte, el Local Development Research Institute (LDRI) promueve a través de la African Open Data Network (AODN) [un programa de open data fellowship en cuatro países africanos](#) (Kenia, Etiopía, Sierra Leona y Ruanda). El objetivo es dar soporte a las distintas agencias, departamentos y ministerios para que sean capaces de desarrollar todo su potencial en áreas como la agricultura, sanidad, educación, contratación o finanzas.

2.5 – COOPERATIVAS Y COMUNIDADES DE DATOS

Las **estructuras cooperativas** pueden aplicarse también al mundo de los datos en un intento de *reequilibrar* la asimetría generalmente existente entre los titulares de los datos y aquellos que buscan utilizarlos en sus servicios y productos. Estas cooperativas no son más que **almacenes de datos**, generalmente en torno a una

temática particular, que se rigen por ciertas **normas** similares a las de las cooperativas tradicionales:



La propiedad es compartida por todos los miembros, por lo que la rendición de cuentas es también más amplia.



Representan a todos sus miembros, **defendiendo sus intereses por encima de todo** y buscando siempre crear valor para ellos.



Potencian el **poder de negociación** del colectivo de sus miembros frente a sus posibles posiciones individuales, por lo que son una herramienta útil cuando se desea generar algún tipo de influencia o cambio.



Sirven también como **herramienta de aprendizaje y empoderamiento** sobre el uso de los datos, tanto propio como por parte de terceras personas.

Por ejemplo, [en los Estados Unidos casi 100 millones de personas son miembros de cooperativas de ahorro y crédito sin fines de lucro](#) (denominadas *credit unions*). Estas instituciones son propiedad de sus miembros y se ocupan de gestionar de forma segura sus datos financieros a la hora de representarlos ante terceros en una amplia variedad de transacciones, como por ejemplo inversiones o seguros.

Otro buen ejemplo es el de [MIDATA.coop](#), una cooperativa de origen suizo que facilita a las personas el control de sus datos médicos. Los miembros de la cooperativa pueden recopilar una variedad de información relacionada con su salud – como por ejemplo registros médicos hospitalarios o los datos recogidos por sus relojes deportivos – y encriptarlos y almacenarlos de forma segura para después compartirlos con quien se desee: su médico personal, sus amigos y familiares o con ensayos e investigaciones clínicas en las que sus datos puedan resultar útiles.

2.6 – RETOS, CONCURSOS Y COMPETICIONES

No siempre es fácil mostrar el impacto de los datos abiertos. Cuando organizamos algún tipo de competición, ya sea ofreciendo algún **incentivo** en forma de premio o no, lo que estamos haciendo en realidad es establecer un **proceso de innovación abierta** que nos ayudará a entender cómo se pueden usar esos datos y para qué.

Este tipo de acciones serán particularmente útiles cuando nos encontremos en un área de conocimiento en la que existen oportunidades, pero está todavía poco explorada y que además generalmente requiere de la participación de distintos agentes para poder tener una visión completa. En ese entorno es poco probable que alguien quiera correr riesgos desarrollando sus ideas cuando todavía no existe un mercado probado para ellas. Es ahí precisamente cuando este tipo de modelo de colaboración puede resultar de especial utilidad. Otras de las ventajas de este modelo son:



El foco está puesto sobre las personas desde el inicio del proceso, ya que se parte de una necesidad real a la que hay que dar una solución apoyándose en el uso adecuado de los datos disponibles.



Se crean **entornos dinámicos e innovadores** en los que las ideas son puestas a prueba y evolucionadas constantemente, utilizando enfoques y metodologías ágiles propias de [la filosofía Lean Startup](#).



Se fomenta no solo la competición sino también la **colaboración**, al reunir a personas y grupos con intereses comunes y perfiles que se complementan en un ambiente altamente interactivo en el que se crean nuevas relaciones y contactos con gran facilidad.

Existen cientos de ejemplos de concursos o [hackathons](#), pero también hay otros modelos más evolucionados y formales como por ejemplo el [Desafío Aporta](#), organizado por la iniciativa nacional de datos abiertos [datos.gob.es](#), o el [Open Data](#)

[Challenge Series](#) llevado a cabo por Nesta y el Open Data Institute (ODI) para explorar soluciones innovadoras en áreas tan dispares y relevantes como trabajo, herencia cultural, alimentación, vivienda, educación, energía o criminalidad.

Otra variante interesante es la que desarrolla la plataforma *Kaggle* – la comunidad online de referencia para los científicos de datos – a través de sus [competiciones](#) para resolver retos en el campo de las ciencias de datos, y gracias a las cuales se han publicado ya [más de 20.000 conjuntos de datos](#).

2.7 – INCUBADORAS Y ACELERADORAS

El principal objetivo de las incubadoras y aceleradoras es **asesorar y apoyar a las ideas o proyectos de negocio con los que trabajan**, contando con diferentes recursos para impulsarlos y conseguir que lleguen a ser rentables y capaces de trabajar de forma independiente en el mercado. Sin embargo, estos modelos de incubación y aceleración han ido evolucionando y especializándose, existiendo hoy en día [incubadoras y aceleradoras centradas específicamente en los negocios basados en los datos y las tecnologías asociadas](#) – tanto en el sector privado, como por ejemplo [Data Elite](#), como en el público como [ODINE](#), [Data Market Services](#) o [Finodex](#).

Además, están surgiendo también nuevos modelos aún más sofisticados y focalizados hacia el fomento de la innovación mediante el uso de los datos, como por ejemplo los dirigidos a la *mentorización* de nuevos científicos de datos como en el caso de [The Data Incubator](#) u otros especializados en la colaboración público-privada como la [European Data Incubator](#) y el [Data Pitch](#) – ambos financiados por la Unión Europea – dando lugar a decenas de [colaboraciones exitosas](#) y [nuevas empresas de datos](#) a lo largo de toda Europa hasta el momento.

Estos últimos programas avanzados de innovación colaborativa y abierta trabajan reuniendo a organizaciones empresariales y del sector público que disponen de datos que desean explotar de diferentes formas con *startups* y PYMEs que puedan trabajar con ellos a través de acuerdos de explotación específicos. El modelo se centra en describir claramente los retos planteados por las organizaciones que suministran los

datos y ayudar a las PYMEs a desarrollar soluciones para hacer frente a estos retos a través de un programa de aceleración en el que además recibirán asesoramiento y un impulso financiero. El objetivo final es que las soluciones desarrolladas puedan llegar a convertirse en negocios sólidos y sostenibles con mayor facilidad gracias a que han nacido partiendo ya de la base a una necesidad real de un cliente real.

CONCLUSIONES

Como hemos visto a lo largo de la primera parte de este informe, los datos abiertos (en mayor o menor medida) ya no pueden considerarse una tendencia exclusiva de las administraciones públicas, sino que **cada vez son más los actores que están compartiendo también sus propios datos**, aumentando así la disponibilidad de información y, por ende, el valor de la propia comunidad de los datos abiertos. **Todos ellos comparten un mismo objetivo, aunque las motivaciones pueden llegar a ser muy diferentes** – imperativos legales, mejora de la imagen, beneficio económico, rendición de cuentas, responsabilidad social, experimentación e innovación, gestión del conocimiento, desarrollo global, protección de la privacidad... cada actor persigue sus propios objetivos, a veces opuestos, a veces compartidos.

Por otro lado, muchas de las iniciativas de apertura de datos gubernamentales que iniciaron su andadura años atrás como pioneras han alcanzado un mayor grado de madurez en el presente y se encuentran ahora con nuevos retos a la hora de escalar sus propias iniciativas. La gestión del ciclo de vida de los datos, los aspectos técnicos y legales, la adaptación de las capacidades de los empleados públicos o la creciente preocupación frente a los potenciales problemas de privacidad son algunas de las barreras con las que las administraciones públicas se tienen que enfrentar cuando intentan llevar sus proyectos de datos abiertos al próximo nivel.

No es de extrañar por tanto que, en vista al creciente número de actores activos en este espacio y en respuesta a las crecientes dificultades a la hora de dar salida y compartir una cada vez mayor cantidad de datos, comiencen ahora a surgir **nuevos modelos colaborativos para una gestión y apertura de los datos más eficaz y efectiva**, tal y como hemos repasado en la segunda parte del informe. Una tendencia que con toda probabilidad seguirá creciendo en un futuro próximo.

Desde la visión experimental de los retos y competiciones, al objetivo empresarial de las incubadoras. Desde el punto de vista de la responsabilidad legal de los *data trust*, al modelo participativo de los *data collaboratives*. Desde la perspectiva formativa de los *data fellowships*, al empoderamiento de las cooperativas de datos. Todos los

modelos existentes en la actualidad comparten algunas similitudes – y en cierta medida también se entremezclan entre ellos dando lugar a otros nuevos modelos – pero cada uno de ellos cuenta al mismo tiempo con su propia esencia, su señal de identidad y su razón de ser.

Y si hay algo que también todos ellos comparten es su carácter experimental. Todos y cada uno de los modelos anteriores surgen como respuesta a distintos intentos de dar solución a los retos en la publicación que antes mencionamos e intentan buscar su hueco en el ecosistema de los datos, a veces promocionados por una organización o entidad en particular y otras veces de manera más espontánea. No obstante, es muy poco probable que podamos encontrar un modelo único de colaboración que pueda cubrir todas nuestras necesidades y sirva para cualquier posible escenario presente y futuro. Así pues, dada la situación actual, lo que cabría esperar en el corto plazo es una todavía mayor proliferación de modelos hasta contar con la experiencia suficiente que permita definirlos y acotarlos mejor. A partir de ahí se pasará con toda probabilidad a una fase de consolidación con algunos modelos de referencia más probados y refinados de forma que puedan ser adoptados más fácilmente.

REFERENCIAS

- MIT Technology review – It is time for a Bill of Data Rights.
<https://www.technologyreview.com/s/612588/its-time-for-a-bill-of-data-rights/>
- BOE – Ley 11/2018, de 28 de diciembre, por la que se modifica el Código de Comercio, el texto refundido de la Ley de Sociedades de Capital aprobado por el Real Decreto Legislativo 1/2010, de 2 de julio, y la Ley 22/2015, de 20 de julio, de Auditoría de Cuentas, en materia de información no financiera y diversidad. <https://www.boe.es/buscar/doc.php?id=BOE-A-2018-17989>
- Fundación COTEC – Guía para la apertura y compartición de datos en el entorno empresarial.
https://cotec.es/media/Guiadatosabiertosenelsectorempresarial_COTEC.pdf
- GIZ – Data for Development: What’s next?
http://webfoundation.org/docs/2018/01/Final_Data-for-development_Whats-next_Studie_EN.pdf
- Observatorio de responsabilidad social corporativa – La Responsabilidad Social Corporativa en las memorias anuales de las empresas del IBEX 35.
<https://observatoriorsc.org/la-responsabilidad-social-corporativa-en-las-memorias-anuales-de-las-empresas-del-ibex-35/>
- Fundación Compromiso y Transparencia – Contribución y transparencia fiscal 2018. <https://www.compromisoytransparencia.com/tematicas-informes/responsabilidad-fiscal>
- UNGP – Data Philanthropy: Public & Private sector data sharing for global resilience. <https://www.unglobalpulse.org/blog/data-philanthropy-public-private-sector-data-sharing-global-resilience>
- Comisión Europea – Directrices para el intercambio de datos del sector privado. <https://eur-lex.europa.eu/legal-content/ES/TXT/HTML/?uri=CELEX:52018SC0125&from=EN>
- ICO – Data sharing code of practice. https://ico.org.uk/media/for-organisations/documents/1068/data_sharing_code_of_practice.pdf

- MIT Connection Science – Data Cooperatives: Towards a Foundation for Decentralized Personal Data Management.
<https://arxiv.org/pdf/1905.08819.pdf>
- Cambridge University Judge Business School for the European Commission - Study on Personal Data Stores. <https://ec.europa.eu/digital-single-market/en/news/study-personal-data-stores-conducted-cambridge-university-judge-business-school>
- Nesta – The Open Data Challenge Series Handbook.
<https://www.nesta.org.uk/documents/424/nesta-odi-odcs-handbook.pdf>
- IEEE – The Personal Data Store Approach to Personal Data Security.
<https://ieeexplore.ieee.org/document/6336755>
- Urban Institute – Data Philanthropy Unlocking the Power of Private Data for Public Good.
https://www.urban.org/sites/default/files/publication/98810/data_philanthropy_unlocking_the_power_of_private_data_for_public_good_2.pdf
- Accenture – Value of data: The dawn of the data marketplace.
<https://www.accenture.com/us-en/insights/high-tech/dawn-of-data-marketplace>
- Gartner – Understand the Data Brokerage Market Before Choosing a Provider.
<https://www.gartner.com/en/documents/3746017/understand-the-data-brokerage-market-before-choosing-a-p>
- Federal Trade Commission – A Call for Transparency and Accountability.
<https://www.ftc.gov/system/files/documents/reports/data-brokers-call-transparency-accountability-report-federal-trade-commission-may-2014/140527databrokerreport.pdf>
- Open Data Institute – Data trusts: lessons from three pilots.
<https://docs.google.com/document/d/118RqyUAWP3WlYyCO4iLUT3oOobnYJGibEhspr2v87jg/edit>
- Data Commons Blueprint – A high-trust, lower-cost alternative to enabling data integration and reuse.
<https://static1.squarespace.com/static/57ad3087579fb3721152748f/t/58aea6b217bffc1271903b2e/1487840951842/Data+Commons+Blueprint+V1.3.pdf>

- GovLab, UNICEF and Omydiar Network – An introduction to data collaboratives. <http://datacollaboratives.org/static/files/data-collaboratives-intro.pdf>
- Peter Suber – Open Access Overview. <http://legacy.earlham.edu/~peters/fos/overview.htm>
- Creative Commons – Open Access. <https://creativecommons.org/about/program-areas/open-access/>
- H2020 Programme – Guidelines to the Rules on Open Access to Scientific Publications and Open Access to Research Data. https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-pilot-guide_en.pdf