



COMO FACILITAR LA COLABORACIÓN EN TORNO A LOS DATOS PARA RESOLVER PROBLEMAS DE LA SOCIEDAD

Datos para crear valor público

Mayo 2022

CONTENIDOS

_Toc104464394

1. INTRODUCCIÓN.....	4
1.1 EL SECTOR PÚBLICO.....	4
1.2 EL SECTOR PRIVADO.....	5
1.3 LA CIUDADANÍA.....	6
2. ¿CÓMO SE INCENTIVA LA COLABORACIÓN?.....	7
2.1 PRINCIPALES BARRERAS.....	8
2.2 BENEFICIOS DE LA COLABORACIÓN.....	9
2.3 INCENTIVOS A LA COLABORACIÓN.....	10
3. ¿CÓMO SE HABILITA LA GOBERNANZA DE LA COLABORACIÓN?.....	12
3.1 GESTIONANDO DATOS PERSONALES APORTADOS POR CIUDADANOS.....	12
3.2 CREANDO NUEVOS CONJUNTOS DE DATOS CON FUERZA DE TRABAJO VOLUNTARIO.....	15
3.3 COLABORANDO ENTRE ACTORES DE DIFERENTES SECTORES.....	17
3.4 INTERCAMBIANDO DATOS DE EMPRESA A GOBIERNO O INSTITUCIÓN INTERNACIONAL (B2G DATA-SHARING).....	20
4. RECURSOS PARA EL APRENDIZAJE.....	23
4.1 REDES DE COLABORACIÓN.....	23
4.2 HERRAMIENTAS Y MARCOS.....	24
4.3 CURSOS Y FORMACIONES.....	26
5. CONCLUSIONES.....	27
6. BIBLIOGRAFÍA.....	28

**Contenido elaborado por José Luis Marín,
Senior Consultant in Data, Strategy,
Innovation & Digitalization.**

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.

Aviso legal: Esta obra está sujeta a una licencia Atribución 4.0 de Creative Commons (CC BY 4.0). Está permitida su reproducción, distribución, comunicación pública y transformación para generar una obra derivada, sin ninguna restricción, siempre que se cite al titular de los derechos (Ministerio de Asuntos Económicos y Transformación Digital a través de la Entidad Pública Empresarial Red.es). La licencia completa se puede consultar en:

<https://creativecommons.org/licenses/by/4.0/>

1. INTRODUCCIÓN

En este informe repasamos algunos de las claves para impulsar e incentivar la creación de **mecanismos de colaboración entre los tres actores principales de estas iniciativas: el sector público, el sector privado y la propia ciudadanía.**

Resolver los complejos desafíos actuales, desde abordar el cambio climático hasta las crisis de salud pública, requiere una inteligente combinación de datos del sector público, de datos de múltiples compañías del sector privado y de datos generados por comunidades de voluntarios.

En los últimos años se han puesto en marcha numerosas **iniciativas basadas en [colaborar en torno a los datos](#)** que han recibido diversos nombres como: “data collaboratives” (colaboraciones de datos), “data institutions” (organizaciones gestoras de datos) u “organizaciones de gestión de datos con fines altruistas reconocidas en la Unión”, pero todas comparten el objetivo de convertirse en iniciativas de colaboración intersectoriales (y públicas/privadas) destinadas a la recopilación, intercambio o procesamiento de datos con el fin de abordar un desafío social.

1.1 EL SECTOR PÚBLICO

El sector público en España, como en el resto de los países de nuestro entorno, lleva más de una década **fuertemente comprometido con el avance hacia la apertura de los datos** que custodia o produce en el ejercicio de sus funciones, con el fin de que puedan reutilizarse para crear valor económico, social, cultural o de cualquier otro tipo.

El foco en esta etapa se ha dirigido principalmente hacia las oportunidades y desafíos de la apertura de datos y de la propia calidad de los datos, bajo la asunción de que, una vez que los datos son reutilizados, se genera valor en diferentes formas, como la mejora de la transparencia o la toma de decisiones más precisa.

Sin embargo, no todos los conjuntos de datos relevantes para la solución de un determinado desafío social se publican como datos abiertos. Por ejemplo, es posible que los conjuntos de **datos relevantes para un problema solo se compartan con un grupo u organización específicos**, situación que se denomina en ocasiones como [datos semiabiertos](#).

La mayoría de los problemas de la sociedad son complejos y la combinación de diversos conjuntos de datos de actores más allá del sector público suele proporcionar, en general, una imagen más precisa y completa de la magnitud y complejidad de estos desafíos. Es esta colaboración la que permite ir más allá de las metas, capacidades y lógicas inmediatas de cada uno de los actores participantes y, por supuesto, **más allá de lo que el sector público puede lograr solo con sus recursos.**

1.2 EL SECTOR PRIVADO

Por otra parte, las clasificaciones de las compañías tecnológicas más valiosas del mundo han evolucionado durante la última década hasta acabar dominadas casi por completo por empresas que han sabido construir sus ventajas competitivas sobre la recolección y explotación de ingentes cantidades de datos. Además, muchas otras compañías de otros sectores de la economía más tradicionales, como la fabricación industrial, el turismo o la distribución minorista, llevan recopilando datos durante años y también han comenzado a identificar su valor.

El acceso a estas grandes cantidades de **datos en manos del sector privado ofrece muchas oportunidades para mejorar la eficiencia, la innovación y la prosperidad económica para la ciudadanía en general.**

[Algunas compañías globales](#) ya han identificado que la colaboración basada en datos por el bien social es una línea importante para ellas y tienen abiertas importantes iniciativas que son el instrumento desde el que gestionan la participación en proyectos colaborativos por el bien común. Algunos ejemplos, con diferentes enfoques y alcance, son:

Luca Data for Good de Telefónica



La misión de LUCA, ahora como parte de Telefonía Tech es **ofrecer la experiencia y las capacidades adquiridas**, a través del proceso de transformación de Telefónica, al resto de organizaciones públicas y privadas. [Telefónica decidió en 2017](#) que el big data y la Inteligencia Artificial podían ser de gran utilidad para el desarrollo de la sociedad y, por ello, **utilizan los datos de conectividad de Telefónica junto con otras fuentes de datos externos** para devolver al mundo el valor de los datos y contribuir a los [objetivos de desarrollo sostenible \(ODS\)](#) de la ONU para 2030. Se han establecido acuerdos con algunas de las organizaciones y ONG más relevantes en el mundo para crear proyectos que contribuyan al bienestar social de las personas más necesitadas.

Open Data Campaign de Microsoft



La campaña de datos abiertos de Microsoft pretende contribuir a **cerrar la "brecha de datos"** y ayudar a las organizaciones de todos los tamaños a darse cuenta de los beneficios de los datos y las nuevas tecnologías que impulsan. La iniciativa se funda desde el convencimiento de que **todos pueden beneficiarse de abrir, compartir y colaborar en torno a los datos** para tomar mejores decisiones, mejorar la eficiencia e incluso ayudar a abordar algunos de los desafíos sociales más apremiantes del mundo.

Data for Good de Facebook



Desde Facebook también se cree que los datos en tiempo real pueden mejorar la respuesta a las crisis del mundo real. Por ello desde esta iniciativa pretenden **aprovechar las estadísticas sobre la comunidad global** que forman los tres mil millones de personas que usan los servicios de Meta **para ayudar a las organizaciones que contribuyen a resolver los problemas sociales**. Desde Data For Good se elaboran mapas, encuestas y estadísticas con el conocimiento de especialistas en temas humanitarios que ahora forman parte del equipo de Data For Good, pero que comenzaron sus carreras en lugares como Cruz Roja, el Banco Mundial, USAID o el Cuerpo de Paz de EE. UU.

A pesar de estos notables ejemplos, a día de hoy no es habitual que los datos de grandes compañías privadas acaben estando a disposición de los proyectos o iniciativas que tratan de resolver problemas de la sociedad. Existen muchas barreras que dificultan la colaboración del sector privado con recursos que con frecuencia son la base de sus propias ventajas competitivas.

1.3 LA CIUDADANÍA

Finalmente, la ciudadanía también se constituye en un actor cada vez más relevante, tanto por el valor que adquieren sus propios datos personales como por su capacidad para generar activos de datos participando en iniciativas que requieren gran cantidad de trabajo voluntario para ser viables.

Los **datos generados en los espacios públicos han sido reconocidos como bien público** por el Foro de la Sociedad del Conocimiento. Las ciudades de Barcelona, Edimburgo, Eindhoven, Gante y Zaragoza lideraron en 2019 un grupo de 81 ciudades para ser más responsables socialmente mediante el uso de datos y desarrollaron los [principios de "datos ciudadanos"](#). Estos principios tienen el objetivo de servir como directrices para las ciudades en sus actividades y contribuir a dar forma al futuro de las políticas urbanas de la UE para ciudades más inteligentes, competitivas, habitables y democráticas.

A su vez, una [consulta pública de la Unión Europea de 2017](#), mostró que existe un apoyo general de los ciudadanos a las acciones que permiten al público tener acceso y compartir datos de salud a través de las fronteras de sus países. Alrededor del 93% de las respuestas estaba de acuerdo en que "el público debería poder gestionar sus propios datos de salud", mientras que el 81% de los participantes creía que "el intercambio de datos de salud podría ser beneficioso para mejorar el tratamiento, el diagnóstico y la prevención de enfermedades en toda la UE".

Aun así, muchas personas desconocen no solo la cantidad y el tipo de datos que producen (por ejemplo, registros de llamadas, datos de servicios digitales, datos en vehículos, registros de atención médica, datos de su actividad deportiva), sino que también ignoran que tienen ciertos derechos sobre sus datos. Por ello, para que la ciudadanía esté dispuesta a contribuir y facilitar el desarrollo de iniciativas de colaboración en la que sus datos sean parte del proceso, es muy importante que sea **plenamente consciente de los beneficios potenciales**, pero también que tenga todas las **garantías sobre el uso adecuado de sus datos**.

En definitiva, bien sea para resolver grandes o pequeños desafíos o para crear valor combinando las capacidades de sector público, sector privado y ciudadanía, **son necesarias plataformas que habiliten, coordinen e incentiven el esfuerzo de las partes**. Por ello, la [agenda España Digital 2025](#) prevé la necesidad de favorecer un fuerte compromiso del sector público para continuar incrementando la apertura de los datos públicos, así como la **creación de mecanismos de colaboración fuerte entre el sector público y el privado**, y el desarrollo de proyectos para usar los datos, tanto públicos como privados o aportados por la ciudadanía, para el bien común.

A lo largo del presente informe vamos a recorrer las diferentes dimensiones del gran desafío que supone **conseguir que los datos que están en manos de actores con muy diferentes motivaciones puedan utilizarse de forma conjunta para resolver problemas de la sociedad**. Para ello, revisaremos fórmulas de gobernanza, examinaremos posibles barreras a la colaboración, describiremos herramientas y formaciones disponibles, y analizaremos una serie de experiencias exitosas. Todo ello con el objetivo de encontrar claves para impulsar e incentivar la creación de mecanismos de colaboración entre los tres actores principales de estas iniciativas: el sector público, el sector privado y la propia ciudadanía.

“El gran desafío es conseguir que los datos que están en manos de actores con muy diferentes motivaciones puedan utilizarse de forma conjunta para resolver problemas de la sociedad”

2. ¿CÓMO SE INCENTIVA LA COLABORACIÓN?

Tradicionalmente las compañías privadas han sido reacias tanto a compartir datos con los gobiernos como a liberarlos para el público en general. Los motivos de estas reticencias, en general, giran siempre en torno al posible impacto negativo que podría tener en sus ingresos la cesión de estos datos. Los datos son activos que se consideran cada vez más valiosos por industrias de todos los sectores y cederlos puede servir para exponer ventajas competitivas o para que otros las construyan basándose en los activos liberados. A pesar de ello, cada vez vemos más compañías que son capaces de explicar a sus accionistas por qué estas colaboraciones son buenas y los beneficios que aportan a la propia compañía y a la sociedad en general.

Por su parte, la ciudadanía cada vez es más consciente del valor de sus datos, así como de los peligros de exponerlos y compartirlos sin las adecuadas garantías respecto a su custodia y utilización de forma agregada.

Por ello **para incentivar la participación es necesario explicar los beneficios que pueden obtenerse tanto de forma individual como colectiva**, a la vez que se garantiza el respeto a la privacidad y la propiedad.

En este contexto el sector público ha generado en los últimos años una serie de nuevas **regulaciones que pretenden generar la confianza necesaria, de forma que pueda crearse valor con la necesaria**

garantía de los derechos de todas las partes. En algunos casos se está también trabajando en el diseño de políticas que persiguen incentivar esta colaboración de forma activa.

A continuación, se exponen cuáles son las principales barreras, beneficios e incentivos para impulsar la colaboración en torno a los datos.

2.1 Principales barreras

Antes de diseñar políticas que actúen como mecanismo de compensación o incentivo que mitigue los riesgos de exponer o ceder activos de datos, lo primero es estudiar y conocer bien cuáles son estos riesgos. Tanto en el caso de las empresas, como de la ciudadanía, podemos citar una serie de barreras que frenan su aportación y participación en iniciativas basadas en datos:

Compañías privadas

- **Pérdida de posición en el mercado:** cuando una compañía de cualquier sector tiene que decidir si expone, aunque solo sea de forma muy parcial, datos que son la base de su negocio, como mínimo analizará con sumo cuidado cualquier paso que vaya a dar. Si existe el riesgo de que, por compartir datos, otros competidores pueden diseñar acciones para erosionar la posición competitiva de quien expone los datos, normalmente las compañías se mantendrán alejadas de este riesgo que no está justificado desde la lógica comercial.
- **Costes de transacción elevados:** incluso aunque se trate de datos con poco valor de negocio para la compañía, la participación en iniciativas de colaboración implica importantes presupuestos solo para la puesta a disposición de los datos. Por ejemplo, la escasa interoperabilidad entre sistemas de información, en la mayoría de los casos exige dedicar importantes inversiones a la hora de ofrecer los datos de forma segura y reutilizable.
- **Poca claridad de los modelos de licencia:** cuando los modelos de licenciamiento no están lo suficientemente claros, las compañías privadas pueden detectar riesgos de pérdida de derechos sobre los datos que comparten. En general, incluso aunque los datos no tengan un gran valor en origen, una compañía normalmente se alejará del riesgo de perder derechos sobre un activo sin ninguna compensación.
- **Cambios regulatorios:** en aquellos sectores que están fuertemente regulados, las compañías podrían ser reacias a exponer no solo los datos sino también información sobre la gestión interna de los datos. Un mayor conocimiento por parte del sector público del funcionamiento interno de estos sectores podría promover cambios regulatorios que modificasen las condiciones de estos mercados. Esto cambios podrían amenazar las condiciones en las que compiten las compañías y por tanto su posición en el mercado.
- **Escasa confianza en los sistemas:** desde una perspectiva técnica, los conjuntos de datos aún no son interoperables y la ausencia de modelos y sistemas dedicados y confiables que permitan el intercambio seguro de datos representa otra importante barrera. Dichos sistemas son esenciales para evitar riesgos al compartir datos (por ejemplo, divulgar información comercial, personal o confidencial)

Ciudadanía

- **Pérdida de privacidad:** la cesión de datos sensibles, como pueden ser los relativos a la salud de los individuos, genera reticencias derivadas de la falta de confianza en el uso que pueda hacerse de los mismos. Que los ciudadanos tengan control sobre el uso que se hace de sus datos es esencial para que la participación sea masiva.
- **Falta de motivación:** en función de cuál sea el problema para el que se pide colaboración, puede resultar difícil movilizar la participación de colectivos que no tienen un interés particular en dicho problema, pero cuya contribución es relevante. Esta motivación es especialmente importante cuando la colaboración requiere algún tipo de esfuerzo o trabajo por parte de los participantes y no solo la cesión de datos personales.
- **Riesgos de ciberseguridad:** la posibilidad de que los datos puedan ser robados y acabar en manos de que quien pueda hacer usos maliciosos es un riesgo percibido por la ciudadanía que se ha puesto de manifiesto incluso en algunas consultas públicas.
- **Desconocimiento en general:** la capacitación digital de una parte de la ciudadanía es una limitación para la participación en determinados tipos de iniciativa, pero no es la única. Las dudas acerca de lo que los marcos regulatorios permiten o no, también genera dudas a la hora de decidir participar.

2.2 Beneficios de la colaboración

Sin embargo, si existen todas estas barreras, ¿qué está motivando a las organizaciones privadas a abrir sus datos en aras del interés público? Analicemos algunos de los principales beneficios para empresas y ciudadanos.

Empresas

- **Responsabilidad social empresarial (RSE) o filantropía de datos,** especialmente de proyectos destinados a países en desarrollo o a acciones que favorecen a grupos desfavorecidos. Muchas compañías, y no solo aquellas de naturaleza más tecnológica, han encontrado en este modo de colaboración una forma de contribución activa y voluntaria a la mejora social.
- **Innovación y aprendizaje:** colaborar con otras entidades en la resolución de problemas permite adquirir nuevos conocimientos, detectar debilidades internas o impulsar la mejora de las propias infraestructuras. Este aprendizaje puede compensar en parte los elevados costes de transacción que tienen estas iniciativas.
- **Mejora de la cadena de valor:** en ocasiones, resolver un determinado problema gracias a la compartición de los datos tiene impactos positivos en la cadena de valor de la compañía y generará beneficios en forma de ahorro de costes o de mejora de la productividad.
- **Intercambio no monetario:** en el marco de la colaboración una compañía privada también puede recibir algunos beneficios en especie a cambio de colaborar cediendo datos.
- **Mejora de imagen o beneficio reputacional:** la transparencia y la generosidad en el uso de los datos privados de la compañía también puede tener como consecuencia un beneficio reputacional que se puede traducir en un aumento de los beneficios.
- **Atracción de talento:** los proyectos de colaboración en torno a los datos, en ocasiones, funcionan para las compañías privadas como fórmulas para la atracción de nuevo talento. En ocasiones, ciudadanos o desarrolladores independientes que aportan trabajo voluntario a estos proyectos acaban siendo contratados por compañías que no los hubiesen conocido en otras circunstancias.

Ciudadanía

- **Interés personal en la resolución del problema concreto**, bien por estar afectados por el mismo o bien por ser sensibles a sus consecuencias.
- **Adquisición de habilidades y aprendizaje**: participar en iniciativas colaborativas permite trabajar nuevas habilidades, generar nuevas relaciones y, en general, aprender sobre un determinado problema.
- **Contribución al progreso de la investigación científica**: aunque se trata de una motivación más difusa en tanto en cuanto los beneficios se obtienen a más largo plazo, una parte de la ciudadanía es sensible a la necesidad de contribuir al progreso de la investigación científica.
- **Satisfacción personal**: la contribución a la resolución de un determinado problema representa una satisfacción de personal que es suficiente incentivo para muchas personas.

2.3 Incentivos a la colaboración

Los beneficios que las compañías o la ciudadanía encuentran en la participación en las iniciativas en las que tienen que aportar sus datos son suficientes solo en algunos casos para que estos proyectos salgan adelante. Existe por tanto mucho margen para incentivar tanto la participación como el liderazgo de nuevas iniciativas de colaboración basadas en datos.

Para impulsar una mayor participación, reduciendo las barreras existentes, una posibilidad es la **creación de mecanismos externos que incentiven dicha colaboración**, ya sean promovidos por las administraciones públicas o por iniciativas privadas. La necesidad de desarrollar incentivos a la colaboración ha sido recogida, por ejemplo, en el [informe final del Grupo de Expertos de Alto Nivel sobre el intercambio de datos de empresa a gobierno \(B2G\)](#),

Algunas de los posibles mecanismos para fomentar esta colaboración están recogidos en las recomendaciones de [este informe](#), que en general podrían ser aplicables a cualquier tipo de colaboración basada en datos y no solo a al intercambio entre gobierno y empresa:

- Establecer **estructuras organizativas** para el intercambio de datos: crear colaboraciones de datos mutuamente beneficiosas entre sector público y sector privado para fomentar su sostenibilidad.
- Desarrollar una serie de esquemas de **incentivos no monetarios** (como incentivos fiscales) a nivel local, nacional y de la UE que puedan beneficiar a las empresas privadas y las organizaciones de la sociedad civil que participen.
- Desarrollar esquemas de **reconocimiento público, como premios, distintivos o marcas**, para socios confiables en el intercambio de datos. Esto aumentaría la reputación positiva de las empresas y organizaciones privadas que ayudan a las autoridades públicas a aumentar el bien de la sociedad.
- La **inversión en el desarrollo de infraestructuras** para compartir por parte del sector público es otra forma de generar estímulos a la colaboración, ya que se reduce el coste de transacción.
- La **inversión directa en investigación**, por ejemplo, en cómo utilizar los datos para el interés público o en el desarrollo de planes para habilitar donación de datos. También se puede investigar cómo obtener más evidencia empírica sobre el impacto cuantitativo del intercambio de datos para el interés público, a través de programas como [Horizonte Europa](#). El [Programa Europa Digital](#) también podría generar incentivos a medio plazo.

Sin embargo, aún nos encontramos en una fase muy inicial ya que muchos de estos incentivos aún deben concretarse y convertirse en medidas que puedan aplicarse en escenarios reales para que sea posible medir sus resultados y evaluar su eficacia. Por ello, proyectos como “[IEDS: incentivos y economía del intercambio de datos](#)”, en el que está trabajando actualmente el Instituto Fraunhofer, permitirán obtener un mayor conocimiento al respecto y el desarrollo de mecanismos sobre cómo incentivar a las organizaciones a participar en colaborativos de datos.

Barreras, beneficios e incentivos para impulsar la colaboración en torno a los datos

CIUDADANOS

EMPRESAS

Barreras

- Pérdida de privacidad
- Falta de motivación
- Riesgos de ciberseguridad
- Desconocimiento en general

- Riesgos de pérdida de posición en el mercado
- Costes de transacción elevados
- Poca claridad de los modelos de licencia
- Cambios regulatorios

Beneficios

- Interés personal
- Adquisición de habilidades y aprendizaje
- Progreso de la investigación científica
- Satisfacción personal

- Responsabilidad social empresarial
- Innovación y aprendizaje
- Mejora de la cadena de valor
- Intercambio no monetario
- Mejora de imagen o beneficio reputacional
- Atracción de talento

INCENTIVOS

Establecer estructuras organizativas para el intercambio de datos

Inversión directa en investigación

Desarrollar una serie de esquemas de incentivos no monetarios

Inversión en el desarrollo de infraestructura para compartir

3. ¿CÓMO SE HABILITA LA GOBERNANZA DE LA COLABORACIÓN?

Para que una iniciativa en la que colaboran varias entidades y particulares tenga éxito y gane adopción, es necesaria alguna forma de gobernanza, esto es, una serie de **procesos y estructuras de toma de decisiones que impliquen a las personas de las organizaciones participantes para llevar a cabo un propósito común**.

Debido a su novedad, aún **no existen análisis sistematizados de las formas ideales de gobernanza de los diferentes tipos de proyectos que intentan contribuir a la solución de problemas sociales** mediante la colaboración basada en datos.

De hecho, aún no tenemos ni tan siquiera un lenguaje común para comunicarnos en torno al fenómeno, y por ello instituciones como Open Data Institute o Govlab están liderando iniciativas para definir y entender mejor la diferentes [fórmulas para acelerar la colaboración en torno a los datos que van surgiendo en el mundo](#).

En la segunda parte del informe **vamos a analizar cuatro de estas fórmulas de gobernanza**, las cuales ya se utilizan con cierta frecuencia para mejorar la colaboración y crear valor. Para cada una de ellas revisaremos casos de éxito que las ponen en práctica.

- Iniciativas que gestionan datos personales aportados por ciudadanos.
- Iniciativas que crean y mantienen conjuntos de datos con fuerza de trabajo voluntario que puede ser de ciudadanos, administraciones públicas o compañías privadas.
- Iniciativas en las que se combinan recursos públicos, privados y de la ciudadanía, y que trabajan en la intersección entre diferentes sectores de actividad.
- Proyectos o iniciativas en las que compañías privadas proveen de datos privados al sector público.

3.1 GESTIONANDO DATOS PERSONALES APORTADOS POR CIUDADANOS

La **protección de datos confidenciales y la concesión de acceso en condiciones restringidas** son la base para la construcción de grandes conjuntos de datos que puedan impulsar la investigación en gran cantidad de áreas, desde la salud pública hasta el urbanismo y la movilidad. La necesidad de crear estos conjuntos de datos surge porque la falta de diversidad en los datos imposibilita encontrar soluciones basadas en modelos que sean representativos de la realidad. Por ello, se fundan iniciativas que proveen y gestionan la infraestructura necesaria para que puedan construirse estos conjuntos de datos a gran escala con aportaciones individuales de la ciudadanía en general.

Además de facilitar una plataforma en la que un individuo pueda realizar y gestionar sus aportaciones de datos privados, bien sean de historiales médicos, de su actividad física, de su historial de movilidad o de cualquier otra área, estas iniciativas diseñan fórmulas para que los colaboradores pueden elegir cómo terceras partes accederán a esos datos y los usos que les autorizan.

3.1.1 Casos de Éxito: Base de datos biomédica a gran escala (UK Biobank)

[UK Biobank¹⁵](#), fundado en 2006, es una iniciativa sin ánimo de lucro para crear y gestionar una **base de datos biomédica** a gran escala y un recurso de investigación que contiene información genética y de salud detallada de medio millón de ciudadanos del Reino Unido. La base de datos se actualiza y amplía regularmente con datos adicionales y es accesible a nivel mundial para investigadores que realizan estudios vitales sobre las enfermedades más comunes y potencialmente mortales.

La visión de la iniciativa es conseguir una gran muestra de individuos con representación de gran cantidad de fenotipos, junto con información genética, de estilo de vida, socioeconómica y de salud médica, y que toda esa información se pueda compartir con los investigadores de todo el mundo de la manera más simple y fluida posible.

Gracias a ella se han podido **realizar investigaciones únicas** como la que se está produciendo actualmente sobre el **control de los niveles de anticuerpos contra el COVID-19** a lo largo del tiempo, lo que contribuye al conocimiento científico de la inmunidad frente a esta enfermedad después de la infección. A través de la gran cantidad de datos que UK Biobank tiene sobre sus participantes, la comunidad de investigación global también puede investigar cómo la genética, el estilo de vida y las condiciones médicas existentes afectan la gravedad de este virus en un grupo sobre otro.

UK Biobank utiliza una **colección de diferentes mecanismos técnicos** para garantizar que los datos de los individuos que custodian y administran sean **utilizados de forma segura** por aquellos a los que se proporciona acceso. El proceso sigue este patrón general:

- Los usuarios, tras un proceso de autorización, pueden descargar conjuntos de datos para los usos aprobados mediante transferencia directa. Los usuarios reciben herramientas para descifrar los conjuntos de datos una vez descargados.
- Todos los conjuntos de datos proporcionados a los investigadores se anonimizan y UK Biobank elimina los identificadores directos e indirectos que podrían permitir la identificación de participantes individuales.
- Los conjuntos de datos proporcionados a los investigadores pueden ser individuales o agregados según sean las necesidades de los investigadores cuyo acceso ha sido aprobado. Los conjuntos de datos que incluyen información sobre participantes individuales están anonimizados.
- UK Biobank se reserva el derecho de realizar una auditoría "para revisar la seguridad, el almacenamiento u otros procedimientos" sobre cualquier información transferida al usuario.



Figura 1: Línea de tiempo del proyecto UK Biobank

La evolución de UK Biobank desde su fundación demuestra el valor de realizar una inversión significativa a largo plazo. Según UK Biobank, [más de 670 grupos de investigación internacionales](#) han accedido a los registros de salud solo durante la pandemia, lo que ha llevado a la publicación de más de 60 artículos científicos de dominio público. En realidad, son ya **más de 2.300 publicaciones científicas las que se han beneficiado de los datos de UK BioBank** y aproximadamente el 75% de las solicitudes de acceso provienen anualmente de fuera del Reino Unido.

3.1.2 Casos de Éxito: Accelerating Medicines Partnership (AMP)

[Accelerating Medicines Partnership \(AMP\)](#) reúne a socios del sector público, múltiples compañías biofarmacéuticas y de ciencias de la vida, organizaciones sin ánimo de lucro y otras organizaciones con el objetivo de **transformar el modelo actual de desarrollo de nuevos diagnósticos y tratamientos**.

Toda la comunidad de investigación biomédica y el público tienen un interés compartido en **reducir los plazos, reducir los costes y aumentar las tasas de éxito del desarrollo de nuevas terapias**. Dada la cantidad y complejidad de los datos, se apuesta por un enfoque sistemático en el que el gobierno, la academia, la industria y los grupos de pacientes trabajen en colaboración.

Se centran en las siguientes enfermedades: enfermedad de Alzheimer, trastornos autoinmunes como la artritis reumatoide y el lupus eritematoso, enfermedades metabólicas comunes, esquizofrenia y enfermedad del Parkinson. Todos los socios de AMP consienten en que **los datos y los análisis que aportan sean de acceso público para la comunidad biomédica** a través de un portal en línea.

Un comité ejecutivo de representantes del Gobierno de Estados Unidos, de las industrias participantes, y de organizaciones sin ánimo de lucro se encargan de supervisar la organización. Los proyectos

individuales se supervisan a través de comités directivos para cada una de las enfermedades en estudio. Cada comité contiene representantes de todos los socios.

La iniciativa combina experiencia pública y privada y datos de diversas fuentes para reducir el tiempo y el coste de desarrollar biomarcadores para el tratamiento de enfermedades. El proyecto intenta superar la fragmentación en la industria farmacéutica y mejorar la innovación en el desarrollo de nuevas terapias con medicamentos. Al combinar datos es más sencillo determinar nuevos objetivos de medicamentos y reducir la repetición inútil de investigaciones que se realizarían varias veces si las empresas siguiesen trabajando en silos.

Aunque en un [artículo publicado en 2019 en Nature](#), en el que se analizan los resultados obtenidos en los primeros años de actividad, se reconoce que aún no se puede atribuir ningún descubrimiento a la iniciativa, sí se demuestra que se han reducido muchos costes asociados a la investigación en numerosos proyectos. Además, se considera que aún es pronto para tener ejemplos concretos debido a los largos periodos que operan en los proyectos de desarrollo de nuevos fármacos.

3.2 CREANDO NUEVOS CONJUNTOS DE DATOS CON FUERZA DE TRABAJO VOLUNTARIO

Crear conjuntos de datos abiertos a los que cualquiera pueda acceder, usar y compartir para promover una misión o causa en particular es una de las formas más complejas de colaboración basadas en datos. Desde que Wikipedia se consolidó como un gran éxito de la colaboración anónima y voluntaria gracias a la posibilidad que ofrecía internet a los individuos para trabajar en un proyecto común si se ponía una plataforma a su disposición, muchos proyectos lo han intentado. Algunos se han consolidado como herramientas globales que utilizamos todos, pero muchos otros no han conseguido sus objetivos y se han quedado en el camino. Algunos factores que se sabe que contribuyen al éxito de estas iniciativas son **la plataforma tecnológica elegida para facilitar la colaboración, la fórmula de gobernanza del proyecto y la capacidad para atraer donaciones iniciales de datos**. Todo ello sin olvidar, claro está, la importancia de tener paciencia para esperar a que millones de pequeñas contribuciones cristalicen en un servicio con un gran valor añadido para la comunidad.

3.2.1 Casos de Éxito: MetaBrainz Foundation

[MetaBrainz Foundation](#) es una organización sin fines de lucro que cree en el acceso libre y gratuito a los datos. Se ha creado para **gestionar bases de datos mantenidas por la comunidad y ponerlas a disposición del dominio público o bajo licencias [Creative Commons](#)**. Sus proyectos incluyen MusicBrainz, la enciclopedia musical, ListenBrainz, AcousticBrainz, BookBrainz y CritiqueBrainz.

Comenzó en 2004 como un **recurso para coleccionistas de música**, pero actualmente ha conseguido establecer **colaboraciones entre voluntarios, artistas y líderes de la industria**. Desde su fundación, se ha convertido en una comunidad global de cientos de miles de editores que colaboran juntos para construir y mantener los conjuntos de datos completos de MetaBrainz.

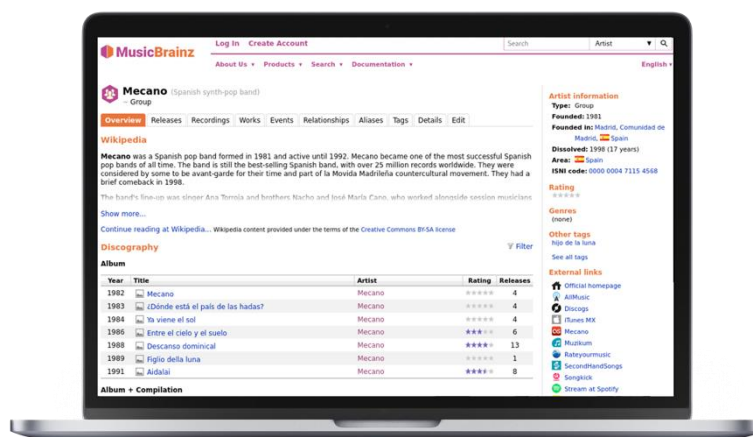


Figura 2: Ejemplo de ficha del grupo musical Mecano en MusicBrainz

Los datos de los diferentes proyectos son recopilados principalmente por voluntarios verificados y se realiza una **revisión por pares para garantizar que sean consistentes y correctos**. El progreso inicial fue lento y la construcción de la infraestructura tomó varios años: siete años para desarrollar la base de datos y cuatro años para que las grandes empresas pensaran que los datos tenían la calidad suficiente para comenzar a usarlos y pagar por ellos. Sin embargo, en la actualidad sus datos son utilizados por organizaciones globales como Amazon, Google, Ticketmaster o la BBC.

La fundación opera con un **modelo de ingresos mixtos** que se basa principalmente en **patrocinadores comerciales y también en donaciones de la comunidad**. Si bien no es obligatorio, MetaBrainz apela a la obligación moral de los usuarios con fines comerciales de contribuir financieramente. Con estos ingresos, la fundación puede continuar con su misión sin la necesidad de otras fuentes de financiación. También se anima a todos los usuarios a que contribuyan al proceso de recopilación de datos para que estos puedan ser lo más completos posible.

3.2.2 Casos de Éxito: eBird

[eBird](#) fue fundado por el [Laboratorio de Ornitología de la Universidad de Cornell \(USA\)](#) en 2002 con una idea sencilla: cada persona tiene conocimientos y experiencias únicas. Ellos decidieron centrarse en las personas a las que les gusta **observar pájaros**. El objetivo del proyecto es recoger estos conocimientos en forma de listas de aves, archivarlas, y compartirlas de forma gratuita con el fin de **habilitar nuevos enfoques basados en datos para la ciencia, la conservación y la educación**.

eBird es el **proyecto de ciencia ciudadana** relacionado con biodiversidad más grande del mundo, con más de 100 millones de registros de aves contribuidos cada año por los *eBirders* de todo el mundo. Es una iniciativa colaborativa de gran éxito dentro de un dominio de conocimiento muy concreto que cuenta con cientos de organizaciones sociales, así como miles de expertos regionales, y cientos de miles de usuarios.

Además, el proyecto desarrolla herramientas para facilitar que la afición de observar pájaros sea más gratificante. Estas herramientas permiten a la comunidad de usuarios el manejo de listas de pájaros,

fotos y grabaciones de audio, así como ver mapas de distribución de especies en tiempo real, y recibir alertas que le permitan saber cuándo se han observado las especies.

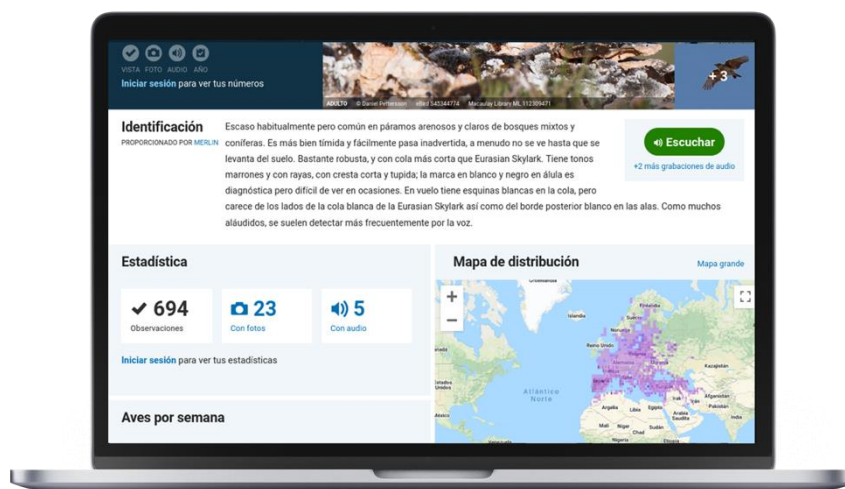


Figura 3: Extracto de la ficha de la Alondra totovía en el proyecto ebird

El equipo de Ciencia eBird utiliza **modelos estadísticos y de aprendizaje automático** para crear visualizaciones y herramientas que ayuden a los responsables de la toma de decisiones, científicos y observadores de aves a comprender mejor la migración, los patrones de abundancia, los límites de distribución y mucho más. El equipo compila datos de eBird sin procesar (el cuándo y dónde las personas reportan aves) e imágenes de satélite de alta resolución de la NASA, NOAA y USGS en modelos estadísticos para predecir cuándo, dónde y en qué números se observan las especies cada semana del año.

Por ejemplo el Laboratorio de Ornitología de Cornell se [asoció con varias organizaciones](#) sin fines de lucro, colaboradores internacionales y agencias para crear [ShorebirdViz](#). Esta herramienta interactiva combina observaciones de aves playeras con modelos estadísticos y de aprendizaje automático para producir estimaciones de abundancia relativa y estimaciones del tamaño de la población de las aves.

eBird juega un **importante y creciente papel en ciencia y conservación**. Las aplicaciones de los datos de eBird van desde la investigación y la monitorización hasta la gestión de las especies o la protección de hábitats. Además, proveen información a los responsables de legislar y diseñar políticas públicas.

3.3 COLABORANDO ENTRE ACTORES DE DIFERENTES SECTORES

La colaboración normalmente produce un mayor valor en la intersección entre disciplinas, pero también es la que presenta unos mayores retos en cuanto a su liderazgo y gobernanza. Cuando hablamos de **colaboración entre organizaciones de diferentes sectores para resolver problemas gracias a los datos**, también encontramos un menor desarrollo que en los otros tipos de proyectos de colaboración que estamos analizando. Existen algunos esfuerzos conjuntos para desarrollar y mantener identificadores, estándares y otras infraestructuras a lo largo de cadenas de valor, pero aún existe un amplio margen para que organizaciones de todo tipo puedan liderar esfuerzos que combinen la contribución de agentes de diferentes sectores.

3.3.1 Casos de Éxito: Mejora de los derechos humanos y el medio ambiente: Open Apparel Registry

El objetivo de [Open Apparel Registry](#) (OAR), fundado en 2019 en Estados Unidos es registrar todas las instalaciones de fabricación de prendas de vestir en todo el mundo y asignarles una identificación única. Abriendo los datos de la cadena de suministro como un bien público y gratuito, esta iniciativa pretende contribuir a la **mejora de los derechos humanos y las condiciones medioambientales dentro y alrededor de fábricas e instalaciones.**

Los datos del Open Apparel Registry son **aportados y utilizados por organizaciones de todo el mundo**, incluidas las principales marcas mundiales, organizaciones de la sociedad civil, organizaciones sectoriales, grupos industriales de fabricación textil, etc. Además de muchos otros beneficios de eficiencia y procesos, la forma en que OAR organiza y presenta los datos en última instancia mejora las vidas de algunas de las personas más vulnerables que trabajan en las cadenas de suministro globales.

La estrategia de la OAR es abrir los datos de la cadena de suministro para el beneficio de todos. Como muchas otras iniciativas de agregación, pretende **transformar datos desordenados e inconsistentes en conjuntos de datos estructurados, disponibles gratuitamente para todas las partes interesadas bajo una licencia de datos abiertos.** Cuando todos los que trabajan en las cadenas de suministro globales disfrutan del mismo acceso a datos de calidad, las oportunidades se abren rápidamente para cambiar la industria hacia un camino más sostenible y equitativo.

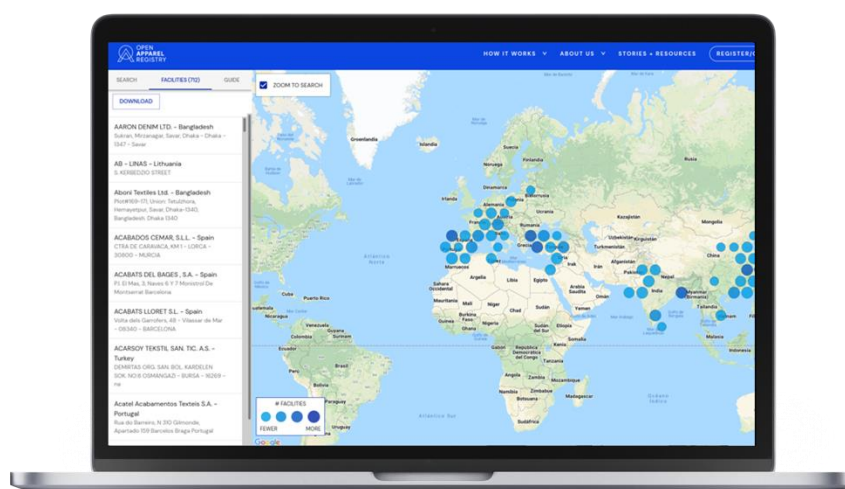


Figura 4: Mapa de instalaciones de Inditex

Si bien la OAR se compromete a **garantizar que cualquier organización pueda acceder a información** de nombres y direcciones, y al identificador asignado por OAR **de forma gratuita, también se ofrecen una serie de servicios de pago** para usuarios con necesidades de datos más sofisticadas. Estas funciones premium son la API de OAR y el mapa embebido, que ayudan a garantizar la sostenibilidad financiera a largo plazo de la herramienta.

Entre las organizaciones que financian la iniciativa se encuentran la fundación [Laudes Foundation](#), dedicada a inspirar actuaciones de la industria por el bien común, la organización sectorial alemana

[The Partnership for Sustainable Textiles](#) de la que son miembros marcas tan destacadas como Adidas o Hugo Boss y la organización filantrópica [Humanity United](#).

Marcas globales como [Clarks utilizan OAR](#) para identificar con precisión dónde están sus proveedores y con qué otras marcas se comparten. También promueve el importante papel de la transparencia dentro de las cadenas de suministro, como palanca para mejorar las condiciones de los trabajadores. Incluso algunas han adoptado el uso del OAR ID como identificador interno en sus propios sistemas de información.

3.3.2 Casos de Éxito: Mejor comprensión de la concesión de subvenciones en Reino Unido: 360 Giving.

La visión de [360Giving](#), fundado en 2014 en Reino Unido, es que la concesión de subvenciones en todos los sectores del país debe estar mejor informada, ser más eficaz y compartir una estrategia. Por ello invitan a las organizaciones financiadoras de subvenciones de todo tipo y tamaño a **compartir información sobre quién, qué y dónde financian, utilizando el estándar de datos creado por 360Giving**. Esto significa que se pueden combinar y comparar miles de millones de euros de datos de subvenciones en todo el Reino Unido, para obtener una imagen más completa de la concesión de subvenciones.

Debido a que son datos abiertos, **cualquiera puede acceder a ellos y usarlos de forma gratuita**. Para que los datos sean más accesibles, se han creado herramientas gratuitas como [GrantNav](#) que facilitan la búsqueda, la descarga y la visualización de los datos.

Los datos están siendo utilizados de diferentes formas según el tipo de organización:

- Las **organizaciones financiadoras** usan los datos y las herramientas para ver cómo hacer el mejor uso posible de los fondos. Esto les ayuda a entenderse mejor a sí mismos y a coordinarse con otros financiadores y no duplicar esfuerzos.
- Las **organizaciones benéficas** destinatarias de los fondos utilizan los datos y las herramientas para conocer mejor a los financiadores y orientar sus esfuerzos de recaudación de fondos.
- Los **planificadores e investigadores** utilizan los datos para obtener una mejor imagen de lo que se financia en todo el Reino Unido.

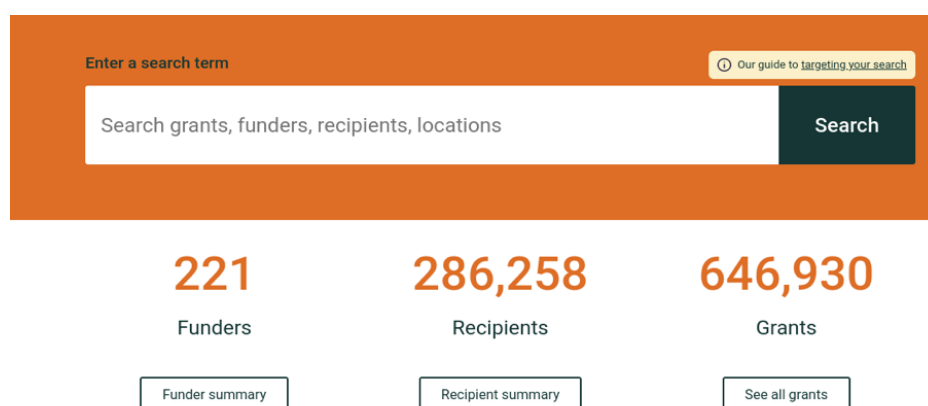


Figura 5: Herramienta para navegar todas las subvenciones agregadas por 360 GrantNav

Entre los datos agregados encontramos ministerios como el de [Justicia](#) o [Defensa](#), ayuntamientos como el de [Birmingham](#) o el de [Londres](#), u organizaciones benéficas de carácter privado como la [Fundación Gatsby](#) o la [Fundación Andrew Lloyd Webber](#).

3.4 INTERCAMBIANDO DATOS DE EMPRESA A GOBIERNO O INSTITUCIÓN INTERNACIONAL (B2G DATA-SHARING).

Los gobiernos necesitan información o datos para, entre otras actividades, tomar decisiones relativas a las políticas que impulsan o la legislación que desarrollan. Cuantos más datos y de mejor calidad sean, mejor informadas podrán ser las decisiones que se tomen. Sin embargo, en muchas áreas de actividad los gobiernos no tienen acceso directo a datos relevantes ya que están en manos del sector privado. Por ello, **el intercambio de datos Empresa-Gobierno (B2G) podría mejorar en gran medida la capacidad de las administraciones para tomar mejores decisiones** políticas en aras del interés público y generar más bienestar social para la sociedad en general.

3.4.1 Casos de Éxito: Análisis de Actividades Pesqueras en la UE

El conocimiento de las actividades pesqueras es un elemento clave para que los responsables políticos planifiquen actividades en el mar (planificación espacial marítima) o evalúen el impacto de la introducción de nuevas áreas marinas protegidas. En este contexto, [MarineTraffic](#) en 2014-2015 colaboró con el Centro Común de Investigación de la Comisión Europea (JRC, BlueHub) al compartir sus datos y experiencia en un gran proyecto de análisis de datos destinado a **comprender mejor las actividades pesqueras en el mar mediante el análisis de grandes cantidades de datos** de seguimiento de embarcaciones para el interés público.

Los datos utilizados en este proyecto de análisis procedían de las redes terrestres de receptores del [Sistema de Identificación Automática](#) (AIS) y contenían información sobre la hora, la posición, la dirección y la velocidad de embarcaciones individuales de más de 15 metros de eslora. De acuerdo con la información que se ha hecho pública sobre el proyecto, para compartir los datos no se empleó ninguna infraestructura en particular, sino que los datos se compartieron directamente entre los equipos involucrados, quienes los combinaron con otras fuentes para su enriquecimiento.

Cada mensaje AIS proporciona la **posición de la embarcación, su velocidad y una marca de tiempo**, en intervalos de cinco minutos. Estos mensajes se clasifican como relacionados con la pesca o con la navegación a través de un **algoritmo de clasificación basado en el análisis de los perfiles de velocidad de las embarcaciones individuales**. Este enfoque de clasificación demostró ser lo suficientemente sólido en el caso de los arrastreros, que representan la mayoría de los buques de más de 15 metros de eslora.

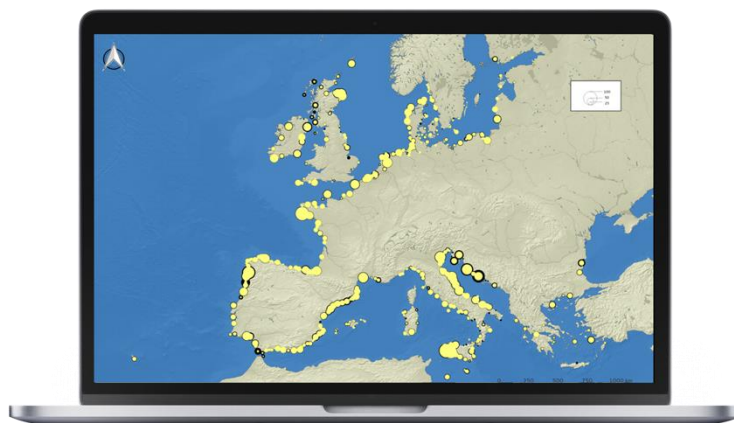


Figura 6: Análisis de Actividades Pesqueras a través de datos del Sistema de Identificación Automática (AIS)

Después de la limpieza y clasificación de datos, los resultados anónimos y agregados, con más de 150 millones de posiciones informadas, se transformaron en **mapas de densidad** y se encuentran disponibles en el [catálogo de datos del Centro Común de Investigación](#).

En este proyecto, MarineTraffic proporcionó los **datos sin coste**, aunque sí estableció algunas **condiciones en la licencia** que afectan a la limitación del propósito para el que pueden utilizarse. El procesamiento de datos históricos AIS, desde septiembre de 2014 hasta septiembre de 2015, dio lugar a la creación **del primer mapa de actividades pesqueras de la UE a escala europea**.

3.4.2 Casos de Éxito: Medición de resiliencia económica ante desastres

[BBVA AI Factory](#) colaboró con [UN Global Pulse](#), la iniciativa de la ONU sobre Big Data e IA para el desarrollo, la acción humanitaria y la paz, en un proyecto que utiliza **datos de transacciones financieras para comprender cómo se comportan las personas antes y después de los desastres naturales**. Para este proyecto, BBVA compartió datos de forma gratuita con la ONU a través de un modelo de preguntas y respuestas (Q&A) en el cual terceros ejecutan algoritmos autorizados a través de la API en la base de datos del propietario de dichos datos para obtener conocimiento. Se trata de un mecanismo que garantiza altos niveles de protección de datos y puede ser utilizado por muchas (terceras) partes, pero que, sin embargo, presenta limitaciones a la hora de enriquecer datos combinando diferentes fuentes.

El proyecto analizó el [impacto del huracán Odile en el estado mexicano de Baja California Sur \(BCS\) en la resiliencia económica de las personas](#). Para ello se analizaron datos de retirada de dinero en efectivo (anonimizados y agregados), de pagos en puntos de venta (POS) de más de 100.000 clientes de BBVA. Los hallazgos mostraron que, a nivel de hogar, las personas gastaron un 50% más de lo habitual en artículos como alimentos y gasolina en preparación para el huracán.

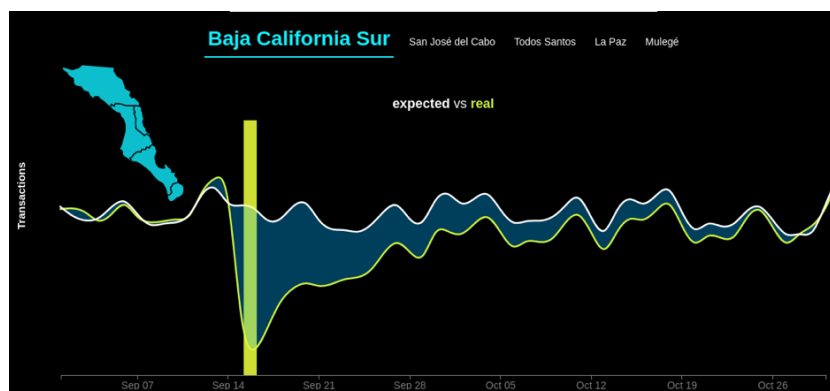


Figura 7: Transacciones esperadas vs transacciones reales

A raíz del huracán Odile, la actividad económica disminuyó en toda la región. Se necesitaron 2 semanas para que las transacciones en puntos de venta y 1 semana en cajeros automáticos recuperaran la normalidad. Al medir el nivel de transacciones exactamente 30 días después del impacto, la investigación mostró que se registraron un 30 % menos de transacciones en puntos de venta y un 12% menos de retiradas de efectivo, en comparación con un período normal.

Tal y como declaró Miguel Luengo-Oroz de UN Global Pulse: *"Este tipo de datos cuantitativos en tiempo real sobre cómo las personas se preparan para un desastre podría usarse para la distribución proactiva y específica de suministros o transferencias de efectivo a las poblaciones más vulnerables"*. Dicho de otro modo, en palabras de Elena Alfaro de BBVA AI Factory *"los nuevos conocimientos pueden ayudar a las autoridades a mejorar la resiliencia de la comunidad, lo que beneficia a los vulnerables y también es bueno para la continuidad del negocio"*.

Se trata de un caso que contribuye a **crear un enfoque replicable y basado en evidencia** para comprender la vulnerabilidad y sienta las bases para el desarrollo de herramientas para ampliar el intercambio de datos B2G. Además, muestra cómo una empresa privada puede mejorar su reputación y potenciar su responsabilidad social empresarial contribuyendo a la consecución de un fin de interés público.

¿A quién aplican los distintos proyectos de colaboración en torno a los datos?	Ciudadanos	Sector privado	Sector público
Gestionando datos personales aportados por ciudadanos	X	X	
Creando nuevos conjuntos de datos con fuerza de trabajo voluntario	X	X	X
Colaborando entre actores de diferentes sectores		X	X
Intercambiando datos de empresa a gobierno o institución internacional (B2G data-sharing)		X	X

4. RECURSOS PARA EL APRENDIZAJE

Como en general la experiencia acumulada es aún muy limitada, se están construyendo **recursos para ayudar a las organizaciones a colaborar de forma exitosa y a reducir las barreras** derivadas de los costes de transacción que supone iniciar una colaboración.

A continuación, hemos recopilado una serie de redes de colaboración, herramientas y metodologías, así como cursos y formaciones que pueden ser de utilidad para aquellas organizaciones que deseen participar en iniciativas de colaboración basadas en datos.

4.1 REDES DE COLABORACIÓN

4.1.1 Red colaborativa de aprendizaje entre pares

Microsoft y ODI lanzaron a finales de 2021 una **Red colaborativa de aprendizaje entre pares que tienen como objetivo que los participantes aprendan de la experiencia y los conocimientos de los demás en sus enfoques para recopilar, mantener y compartir datos**. Está especialmente enfocada a iniciativas de colaboración basadas en datos que se encuentren en sus fases iniciales de desarrollo.

La red se formó con una primera cohorte de seis iniciativas de colaboración basadas en datos que abordan importantes desafíos sociales y económicos en sectores como el transporte y la salud, entre otros. Muchos de los participantes aún se encontraban en una etapa inicial de concepto y participaron en la red para desafiar y validar sus primeros diseños e identificar desafíos clave, así como diferentes formas de abordarlos.

Se han documentado una serie de [aprendizajes a partir de las experiencias de esta primera cohorte](#). Se espera que compartir estas historias ponga de manifiesto algunos de los principales desafíos para las colaboraciones de datos y muestre cómo se puede ayudar a mejorar las posibilidades de éxito con diferentes tipos de apoyo.

La segunda cohorte fue elegida después de un extenso proceso de entrevistas a las candidaturas recibidas en la [convocatoria abierta en agosto de 2021](#), que tuvo como resultado la elección de siete organizaciones con diversos aspectos en común, así como experiencia que ofrecer a los demás en los desafíos planteados.

4.1.2 Red colaborativa de gestores de datos: Data Stewards Network (DSN)

La [Data Stewards Network \(DSN\)](#) es una comunidad de práctica fundada con el apoyo de la Fundación Rockefeller, que tiene como objetivo principal encontrar nuevas formas de reducir los altos costos de transacción que supone el lanzamiento de una iniciativa de colaboración basada en datos.

Los miembros de esta comunidad de práctica pueden aprender de sus pares y compartir sus propios conocimientos y experiencia en administración de datos. Los principales beneficios que obtienen los miembros de esta red son el acceso a pares en múltiples sectores; un medio para desarrollar posiciones y prácticas comunes entre pares; acceso a recursos como al repositorio de memorandos de entendimiento (MOU) y herramientas para desarrollar acuerdos de intercambio de datos; acceso a recursos como marcos de responsabilidad de datos.

La red celebra reuniones periódicas virtuales y presenciales centradas en las solicitudes y necesidades de los miembros y es un excelente escaparate para conocer las oportunidades de colaborar o asociarse con otras empresas y partes interesadas del sector público en la resolución de diferentes problemas a través de los datos.

4.2 HERRAMIENTAS Y MARCOS

4.2.1 El Marco de las 9Rs

El [marco de las 9Rs](#), desarrollado por el Laboratorio de políticas de datos abiertos ([The Open Data Policy Lab](#)) es un **método para describir y validar el *business case* para resolver un problema público a través de la reutilización de datos.**

El Marco 9R consta de nueve motivaciones que han sido identificadas a lo largo de varios años de experiencia en el estudio y establecimiento de colaboraciones de datos, categorizadas por diferentes tipos de retorno de la inversión: licencia para operar, valor de marca o conocimiento e ideas.

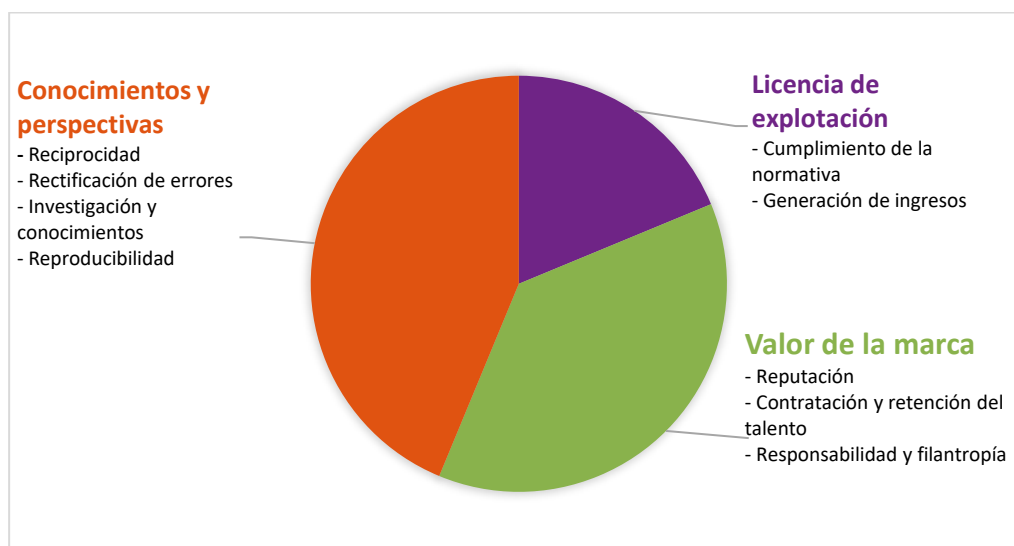


Figura 8: Business case para la reutilización de datos por el bien público

4.2.2 Data Collaboratives Canvas

El [Data Collaboratives Canvas](#) es una herramienta desarrollada por GovLab que **describe 8 fases para diseñar e implementar una colaboración de datos (asociación) a nivel institucional**. Establecer una colaboración de datos efectiva requiere una comprensión detallada de la iniciativa, es decir, una comprensión profunda del problema que se abordará (en particular desde la perspectiva de la demanda) y de la oportunidad que brinda el intercambio de datos entre sectores.

Para cada una de las 8 fases que identifica se proporcionan un número de recursos de ayuda para completarla: cuestionarios, manuales, ejemplos, casos de estudio, etc.

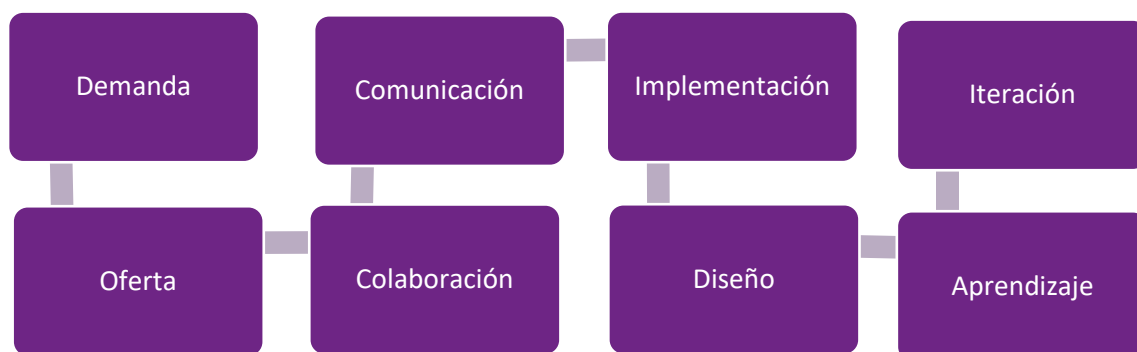
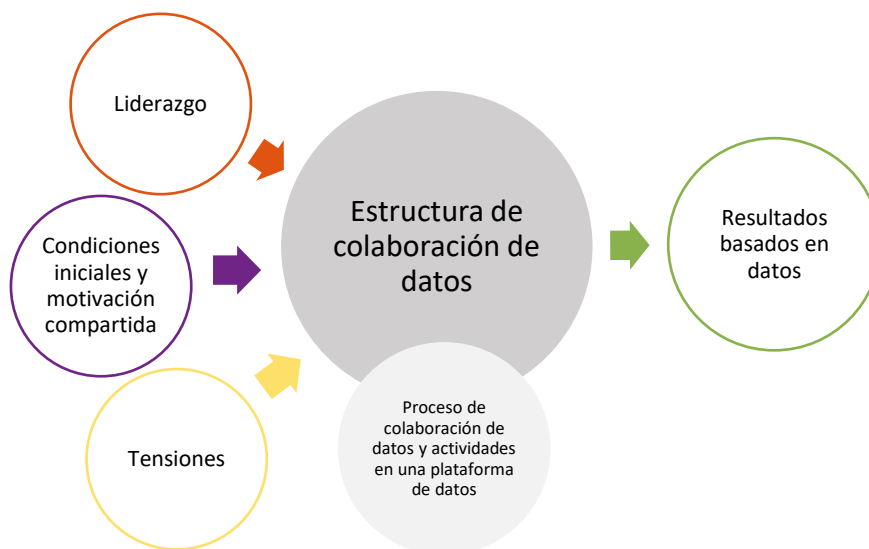


Figura 9: Las ocho fases del Data Collaboratives Canvas

4.2.3 Marco para la Gobernanza de las Colaboraciones de datos

El [marco para la Gobernanza de las Colaboraciones de Datos](#) fue publicado en 2021 como resultado de un estudio que desarrolló un marco teórico y lo validó con la metodología de living lab.

Entorno institucional



De acuerdo con el marco, **el entorno institucional se define como el contexto en el que tiene lugar la colaboración basada en datos**. Las condiciones iniciales y la motivación compartida para participar en la colaboración de datos influyen en la estructura (normas y reglas o prácticas de participación) y en el proceso (confianza, comunicación, comprensión compartida del problema y uso real de los datos), lo que conduce a un resultado de la colaboración basada en datos. La estructura y el proceso trabajan en estrecha colaboración y son difíciles de separar. Dado que la tecnología puede facilitar las actividades

de datos, este componente se incorpora como parte del proceso de colaboración de datos. Finalmente, el liderazgo y las tensiones influyen en la estructura y el proceso.

4.3 CURSOS Y FORMACIONES

4.3.1 Administración de datos: DATA STEWARDSHIP

El programa de formación en línea “[data stewardship](#)” está estructurado en 8 módulos para ayudar a quienes están trabajando en el **desarrollo de una estrategia de reutilización de datos para resolver problemas públicos**. A lo largo de un período de ocho semanas, los participantes adquieren una mayor comprensión de las tres olas de datos abiertos y la capacidad de promover la responsabilidad y la administración de datos dentro de su organización.

Si bien el acceso al programa es gratuito, la participación se limita a pequeñas cohortes que trabajan en problemas de colaboración de datos o actúan como administradores de datos en sus organizaciones.

El programa está coordinado por TheGovLab pero entre sus docentes se encuentran personas con importantes responsabilidades en iniciativas de datos en organizaciones internacionales como UNICEF, Naciones Unidas o la OCDE, compañías globales como MasterCard, Microsoft o Telefónica, gobiernos como el del Reino Unido o la Columbia Británica y miembros de fundaciones como la ISI Foundation, el World Resources Institute o la Knight Foundation.

4.3.2 Resolviendo problemas públicos con datos

El curso “[Resolviendo problemas públicos con datos](#)” examina **cómo se pueden usar los datos para mejorar la toma de decisiones y la resolución de problemas en el sector público**. A través de ejemplos del mundo real y casos de estudio, se discuten los principios fundamentales de la ciencia de datos para ayudar a fomentar una mentalidad analítica de datos. El objetivo es que el estudiante adquiera la capacidad de definir y aprovechar el valor de los datos para lograr su misión pública.

Este curso, desarrollado por [Govlab Academy](#), consta de 4,5 horas video y otros recursos adicionales y no requiere ni asume ninguna experiencia previa en informática o estadística. Todos los [recursos están abiertos](#) y pretende que al finalizar los alumnos puedan articular una propuesta de valor para el uso de datos en la resolución de problemas en el sector público, incluido cómo aprovechar los datos para mejorar la toma de decisiones.

5. CONCLUSIONES

Estamos observando cómo existe un **creciente número de iniciativas para la colaboración basadas en datos** en las que participan sector público, sector privado y la ciudadanía en general, incluyendo las transferencias de datos entre gobierno y empresas, que se basan en la idea de que pueden **proporcionar mejoras en el bienestar socioeconómico** o contribuir a la **solución de desafíos de la sociedad**.

Aunque son necesarios más esfuerzos, en los últimos años se viene trabajando a nivel europeo en una serie de nuevas **regulaciones como la propuesta de [Ley de Gobernanza de Datos](#), la propuesta de [Ley de Datos](#) o la propuesta de [Ley de Inteligencia Artificial](#)** que pretenden generar o mejorar la confianza necesaria para que pueda crearse valor con la necesaria garantía de los derechos de todas las partes. También existe una sensibilidad cada vez mayor hacia la creación de mecanismos externos que incentiven la colaboración sobre todo pensando en reducir las barreras para que el sector privado pueda aportar sus activos de datos y sus conocimientos sin riesgos para sus actividades de negocio.

En este sentido, informes como el realizado por el [Grupo de Expertos de Alto Nivel sobre el intercambio de datos de empresa a gobierno \(B2G\)](#), realizan recomendaciones muy interesantes para que las colaboraciones basadas en datos puedan convertirse en una **práctica responsable, sostenible y escalable en la Unión Europea**. El informe se pide de forma explícita a la Comisión, a los Estados miembros y a todas las partes interesadas que tomen las medidas necesarias, en sus diferentes funciones, para poner a disposición más datos y aumentar su reutilización para el bien común.

Como puede verse a través de los casos de éxito analizados en este informe y en otros muchos como los que están documentados en el Anexo II del informe final del Grupo de Expertos de Alto Nivel sobre el intercambio de datos de empresa a gobierno (B2G), para que una iniciativa en la que colaboren varias entidades y particulares tenga éxito y gane adopción, es necesaria alguna forma de gobernanza, esto es, una serie de **procesos y estructuras de toma de decisiones que implican a las personas de las organizaciones participantes para llevar a cabo un propósito común**. Debido a la escasa experiencia existente, aún **no hay análisis sistematizados** de las formas ideales de gobernanza de los diferentes tipos de proyectos que intentan contribuir a la solución de problemas sociales mediante la colaboración basada en datos.

Lo que parece claro es que para resolver grandes o pequeños desafíos públicos o para crear valor combinando las capacidades de sector público, sector privado y ciudadanía son necesarias plataformas que habiliten, coordinen e incentiven el esfuerzo de las partes.

Finalmente, encontramos que, como respuesta a las dificultades, se está produciendo la formación de redes de colaboración, la publicación de herramientas y metodologías, así como la propuesta de cursos y formaciones que pueden ser de utilidad para aquellas organizaciones que deseen participar en iniciativas de este tipo. Todo ello con el fin de que la poca experiencia existente pueda ayudar a quienes están dando aún los primeros pasos en la resolución de problemas públicos a través de la colaboración basada en datos.

6. Bibliografía

EUROCITIES . EUROCITIES principles on citizen data: 10 data principles for the common good. [en línea] [fecha de consulta: 25 Abril 2022] Disponible en: <https://eurocities.eu/latest/eurocities-principles-on-citizen-data-10-data-principles-for-the-common-good/>

Comisión Europea. Synopsis report of the public consultation on Digital transformation of health and care in the context of the Digital Single Market. [en línea] [fecha de consulta: 25 Abril 2022] Disponible en: <https://digital-strategy.ec.europa.eu/en/synopsis-report-public-consultation-digital-transformation-health-and-care-context-digital-single>

Telefónica. Data as a force for good. [en línea] [fecha de consulta: 25 Abril 2022] Disponible en: <https://www.slideshare.net/LUCA-D3/data-as-a-force-for-good>

Gobierno de España, Ministerio de Asuntos Económicos y Transformación Digital. España Digital 2025. [en línea] [fecha de consulta: 25 Abril 2022] Disponible en: https://portal.mineco.gob.es/ca-es/ministerio/estrategias/Paginas/00_Espana_Digital_2025.aspx

Comisión Europea. Towards a European strategy on business-to-government data sharing for the public interest. [en línea] [fecha de consulta: 25 Abril 2022] Disponible en: <https://op.europa.eu/es/publication-detail/-/publication/d96edc29-70fd-11eb-9ac9-01aa75ed71a1>

Bargh, Mortaza S. and Choenni, Sunil and Meijer, Ronald. Meeting Open Data Halfway: on Semi-Open Data Paradigm [en línea] [fecha de consulta: 25 Abril 2022] Disponible en: <https://dl.acm.org/doi/10.1145/2910019.2910037>

Datos.gob.es. Fórmulas para acelerar la colaboración en torno a los datos. [en línea] [fecha de consulta: 25 Abril 2022] Disponible en: <https://datos.gob.es/es/blog/formulas-para-acelerar-la-colaboracion-en-torno-los-datos>

Nature. Massive NIH–industry project opens portals to target validation. [en línea] [fecha de consulta: 25 Abril 2022] Disponible en: <https://www.nature.com/articles/d41573-019-00033-8>

Wikipedia. Sistema de Identificación Automática. [en línea] [fecha de consulta: 25 Abril 2022] Disponible en: https://es.wikipedia.org/wiki/Sistema_de_Identificaci%C3%B3n_Autom%C3%A1tica

Centro de Investigación Común. Comisión Europea. AIS derived high resolution fishing effort layer for European trawlers of more than 15 meters long 2014-2015. [en línea] [fecha de consulta: 25 Abril 2022] Disponible en: <https://data.jrc.ec.europa.eu/dataset/jrc-fad-ais1415>

The GovLab, Open Data Policy Lab. The “9Rs Framework”: Establishing the Business Case for Data Collaboration and Re-Using Data in the Public Interest. [en línea] [fecha de consulta: 25 Abril 2022] Disponible en: <https://opendatapolicylab.org/articles/business-case-the-9rs-framework-establishing-the-business-case-for-data-collaboration-and-re-using-data-in-the-public-interest/index.html>

GovLab Academy. Solving Public Problems with data. [en línea] [fecha de consulta: 25 Abril 2022] Disponible en: <https://govlabacademy.org/courses/solving-public-problems-with-data.html>

EUR-Lex. Propuesta de reglamento del Parlamento Europeo y del Consejo relativo a la gobernanza europea de datos (Ley de Gobernanza de Datos). [en línea] [fecha de consulta: 25 Abril 2022] Disponible en: <https://eur-lex.europa.eu/legal-content/ES/TXT/HTML/?uri=CELEX:52020PC0767>

EUR-Lex. Propuesta de reglamento del Parlamento Europeo y del Consejo sobre normas armonizadas para un acceso justo a los datos y su utilización (Ley de Datos). [en línea] [fecha de consulta: 25 Abril 2022] Disponible en: <https://eur-lex.europa.eu/legal-content/ES/TXT/HTML/?uri=CELEX:52022PC0068>

EUR-Lex. Propuesta de reglamento del Parlamento Europeo y del Consejo

por el que se establecen normas armonizadas en materia de inteligencia artificial (ley de inteligencia artificial) y se modifican determinados actos legislativos de la unión. [en línea] [fecha de consulta: 25 Abril 2022] Disponible en: <https://eur-lex.europa.eu/legal-content/ES/TXT/HTML/?uri=CELEX:52021PC0206>

