



*Alberto Palomo Lozano\**  
*Manuel Gutiérrez de Diego\*\**

## **BIG DATA WEBS: COLECTIVIZACIÓN DEL VALOR DEL DATO CON BASE EN ECOSISTEMAS FEDERADOS**

Este artículo analiza cómo el desarrollo de una infraestructura común a nivel UE para la interoperabilidad fomenta la colaboración y la compartición de recursos de datos. Desde el diseño, dicha infraestructura articula una innovadora capacidad de interrelación entre los diferentes actores de la nueva Economía del Dato europea (segura, confiable y soberana), favoreciendo el desarrollo de un mercado económico que se estima que crecerá hasta representar el 6% del PIB de la Unión en 2025. Creemos que alrededor de esta infraestructura florecerán los ecosistemas federados de datos, entornos de compartición y explotación que, gracias a una gobernanza participativa, fomentan dinámicas de negocio *win-win* con las que crear valor añadido más allá de la suma de las partes, y cuyo reparto se orquesta de forma equitativa.

**Palabras clave:** datos, *data-driven*, confianza, soberanía digital, compartición de datos, infraestructura digital.

**Clasificación JEL:** H00, H44, O31, O43.

### **1. El importante valor económico del dato**

La Economía del Dato hace referencia a la oferta y demanda global de datos que se generan de forma instantánea a través de aplicaciones informáticas, objetos conectados y de todos los posibles servicios relacionados. En un mundo hiperconectado, cualquier interacción

es susceptible de generar datos. Si queremos estudiar y cuantificar el valor latente en los mismos, la clave está en el potencial interés por estos: un conjunto de datos puede resultar de utilidad para varias empresas que operan en una misma cadena de valor, para competidores de un mismo sector, o para organizaciones de otros sectores con los que se guarde una relación comercial. Incluso es posible que surjan intereses entre empresas que aparentemente no tienen ninguna conexión entre sus modelos de negocio; de hecho, este es probablemente el tipo de impacto más transformador, que abre una puerta a un futuro de oportunidades por (co)desarrollar. ▷

\* Chief Data Officer del Gobierno de España. Secretaría de Estado de Digitalización e Inteligencia Artificial, Ministerio de Asuntos Económicos y Transformación Digital.

\*\* Consultor Sénior de Estrategia. Oficina del Dato, Secretaría de Estado de Digitalización e Inteligencia Artificial, Ministerio de Asuntos Económicos y Transformación Digital.

Versión de julio de 2023.

<https://doi.org/10.32796/bice.2023.3160.7645>

Por supuesto, los actores implicados en la Economía del Dato son múltiples. Por un lado, están las organizaciones que generan los datos, así como las que los utilizan para mejorar sus procesos productivos o para desarrollar nuevos productos y servicios. Por otro lado, están los proveedores de tecnología y servicios que habilitan el almacenamiento, procesamiento y explotación de grandes cantidades de datos. Además, las Administraciones públicas son también una parte importante, tanto en su capacidad de articular procesos de esta misma naturaleza como en lo relativo a una regulación digital que cada vez juega un papel más clave en las relaciones que caracterizan dicha economía. La Economía del Dato afecta a todos los sectores productivos: todas las organizaciones y administraciones generan y utilizan datos en su actividad; si bien es cierto que algunos sectores tienden a utilizar datos de una forma más intensiva (el sector financiero o el de la salud podrían incluirse entre estos). También en algunas industrias las barreras existentes son mayores, incluso culturalmente, y también a lo largo de la dimensión geográfica encontramos diferencias en los niveles de conocimientos técnicos, o en la disponibilidad de recursos para invertir en tecnología.

Según estimaciones, la Economía del Dato europea está destinada a generar un mercado económico que en 2025 alcanzará los 827.000 millones de euros para los 27 Estados miembros —o alrededor del 6% del PIB (European Commission, Directorate-General for Communications Networks, Content and Technology, 2020)—, en gran medida relacionado con el crecimiento del internet de las cosas y la ingente cantidad de datos que estos dispositivos son capaces de originar. Estas estimaciones se basan en diferentes modelos econométricos, por ejemplo, el basado en la venta de datos, o el

que mide la explotación de los datos propios a una organización. También se generan beneficios para los Estados en los que operan dichas entidades, ayudando a estimar la situación económico-financiera de un país, y vertebrando un tejido económico innovador con base en la creación de nuevas empresas y empleos. No obstante, es importante reseñar que esta apuesta por la prosperidad y la innovación debe ser aprovechada sin que ello suponga generar desigualdades, hipotecar el futuro industrial y tecnológico, y sin comprometer los valores y derechos fundamentales alcanzados por la ciudadanía.

Bajo este contexto, la Comisión Europea publica en febrero de 2020 una oportuna, a la par que ambiciosa, [Estrategia Europea de los Datos](#), cuyo objetivo es convertir a la Unión Europea en líder de una sociedad dirigida por los datos. La creación de un mercado único de datos, al que después nos referiremos en mayor profundidad, debe permitir que estos fluyan libremente por la Unión, no solo geográficamente, sino entre sectores, en beneficio de las empresas y organizaciones del tercer sector, los investigadores y las Administraciones públicas.

Este posicionamiento político de carácter estratégico pivota sobre dos ejes tecnológicos actualmente muy en boga: los datos y los servicios *cloud* (o, más genéricamente, el modelo «como Servicio»/«as-a-Service» que ofrece la nube pública). Podemos decir que uno de los objetivos principales de esta Estrategia de los Datos es fomentar el uso del *big data* y la inteligencia artificial en Europa, y para ello se están llevando a cabo iniciativas para mejorar el acceso a los datos y a las tecnologías con que explotarlos. Se ha diseñado un magnánimo plan de trabajo para la transformación digital de empresas y servicios públicos, a lo largo de una serie de iniciativas como: ▷

- el establecimiento de normas claras y justas sobre el acceso y reutilización de conjuntos de datos,
- la inversión en herramientas e infraestructuras de próxima generación para almacenar y tratar los datos,
- la homogeneización y federación de las capacidades europeas de computación en nube, y
- la puesta en común de datos europeos en sectores clave, creando espacios de datos comunes interoperables.

Lo novedoso aquí es que la estrategia no sólo aspira a desarrollar nuevas capacidades y metodologías con las que potenciar las sociedades y economías europeas con base en las dos disciplinas señaladas, sino también a conectarlas entre sí. Y esto está motivado a raíz de la fuerte interrelación que se observa entre el diseño y el despliegue de infraestructuras digitales y su uso innovador aplicadas a sectores concretos. Esta sinergia ha resultado especialmente notable en campos como la inteligencia artificial, cuya efervescencia actual surge gracias a una acompasada coordinación entre la disponibilidad de vastos conjuntos de datos y el uso de tecnologías disruptivas de cómputo paralelo. Esta interrelación entre tecnología y el mundo digital supone filosóficamente la intersección entre la Ley de Moore, que heurísticamente sostiene que el número de transistores en un microprocesador se duplica aproximadamente cada dos años y, por tanto, también se dobla teóricamente la capacidad de cómputo, y la Ley de Metcalfe, que establece que el valor de una red de telecomunicación es proporcional al cuadrado del número de usuarios conectados a ella.

Por ello, y dado que el crecimiento en los datos disponibles coincide con un aumento en la capacidad de procesamiento, podemos hacer

un uso mucho más efectivo de los mismos. Esto es justo lo que ocurrió con el artículo seminal de Krizhevsky *et al.*, (2012), que revolucionó el campo de la visión computacional a base de entrenar una red neuronal profunda con 1,2 millones de imágenes, gracias al uso de dos tarjetas gráficas que las procesaban en paralelo, frente a un modelo secuencial tradicional por el que eso no habría sido posible por el extenso tiempo en que habría incurrido. Un fenómeno similar también pasa, y aunque en este caso la arquitectura informática es diferente, gracias al desarrollo del campo del *big data*, donde se capitaliza en el modelo de «divide y vencerás», partiendo grandes conjuntos de datos en piezas más pequeñas para analizarlas en paralelo, y consolidando luego los resultados de cada minianálisis.

Como vemos, se da, por tanto, una exuberante oportunidad de generar sinergias con base en el acceso, la transformación y el uso de los datos, aumentando estos procesos con capacidades tecnológicas avanzadas como las que ofrece un modelo *as-a-Service* como el de la nube. Esto es algo que las grandes corporaciones tecnológicas han entendido a la perfección, que más allá de proveer de productos o servicios técnicos innovadores, usan también sus plataformas digitales para comercializar ofertas de terceros, dando forma a mercados digitales que intermedian entre proveedores y consumidores de objetos y recursos. Estos foros donde se cita la oferta y la demanda tienen, por supuesto, una traslación natural al ámbito de los datos, y más aún si consideramos la condición no rival de los mismos, que hace que se puedan copiar y almacenar (a un coste cada vez más exiguo) para explotarse en diferentes contextos, sin que eso afecte negativamente al dueño original de sus derechos. ▷

Esta posibilidad de explotación concurrente y no rival de los conjuntos de datos abre nuevas líneas de negocio basadas en la compartición de datos entre organizaciones, a través de las que generar escenarios *win-win* donde el proveedor los reditúa (bien monetizándolos o de otra forma) y el consumidor los explota. Sin embargo, estos escenarios se encuentran todavía limitados a contextos nicho de especialización sectorial, fuertemente condicionados por la confianza entre las partes, y donde esta suele además ligarse a extensos y detallados acuerdos contractuales. En esas condiciones, el desarrollo de un mercado alrededor de los datos se hace más complejo y lo limita severamente. Es por este motivo que la citada Estrategia Europea de los Datos tiene como *leitmotiv* afianzar e impulsar el Mercado Digital Único. Gracias a permitir que los datos fluyan más libremente entre los países y los sectores económicos, se crea un mercado distribuido con el que mejorar la economía digital y el crecimiento económico.

Así, de forma similar al acero y el carbón a partir de 1951, las contrapartes ejecutan en dicho mercado transacciones punto-a-punto, y esto sirve para fomentar la cooperación económica y digitalizar las diferentes cadenas de valor industriales. Bajo este contexto, el dato pasa de ser sólo el subproducto resultante de la interacción de aplicaciones digitales, con utilidad en auditorías y depuración de procesos, a convertirse también en una materia prima que puede reutilizarse de múltiples formas, generando valor añadido incluso a nivel *cross-sectorial*, a partir del uso de tecnologías habilitadoras como el internet de las cosas (IoT), el *big data*, la inteligencia artificial, los gemelos digitales, la realidad virtual / realidad aumentada, o el 5G. Es innegable que el impacto de un mercado único de datos puede ser tremendamente fructífero

para las cadenas de valor económicas, con beneficios potenciales que incluyen la mejora de la eficiencia organizacional, el aumento de la innovación y la creación de nuevos modelos de negocio. Es por ello que la Comisión Europea complementa su estrategia para los datos con otras sectoriales, como por ejemplo el [Plan de Acción para una Economía Circular](#).

## 2. Del modelo de plataforma al de ecosistema

El contexto descrito para el mercado único de datos paneuropeo guarda grandes similitudes con los modelos de plataforma ya en uso en la actualidad, traccionando los dos mismos ingredientes clave: la existencia de un mercado dinámico de oferta y demanda (en este caso de conjuntos y servicios de datos), y contar con un instrumento que actúa como conector entre sus participantes. Así, el modelo económico subyacente en estas plataformas se centra en la creación de valor a través del desarrollo y la gestión de redes de usuarios, al permitir que estos se conecten entre sí, realicen transacciones, compartan recursos o colaboren de alguna manera para el beneficio mutuo. Con base en un espacio o infraestructura digital común, se proporcionan herramientas y servicios que permiten aprovechar las ventajas de la red. Y en cuanto a las formas de repartir el valor, pueden ir desde un modelo de negocio basado en tarifas o comisiones por transacción hasta otros de ingresos por venta de activos o por servicios *premium*.

Sin embargo, un aspecto novedoso de la citada Estrategia de los Datos de la UE es la búsqueda de la colectivización en la captura de valor. Frente a los modelos de plataforma, donde gran parte del valor generado se retiene en ▷

el proceso de intermediación, la UE apuesta por un esquema de ecosistema que, acorde con los principios europeos, evite la generación de dependencias y posiciones de dominancia en los mercados. Este paradigma se sustenta sobre las bases del federalismo, de forma que la clave no reside en una infraestructura única donde todos los participantes desarrollan su actividad, sino en una gobernanza mínima que habilite la interacción entre ellos y que les permita retener su autonomía y poder de decisión en cuanto a las transacciones de datos, en las que tomar o no parte según las condiciones del momento. Es decir, la plataforma deja de ser el elemento central sobre el que todos operan, y son las reglas de gobernanza las que, por el contrario, toman esas riendas.

La fórmula reside, por tanto, en buscar cómo articular un ecosistema más amplio, con base en un modelo federativo donde la interoperabilidad existe desde el diseño, y para el que se busca una activa colaboración entre los diferentes actores en la definición de las «reglas de juego». Esto permite ampliar las oportunidades procedimentales y operativas para los participantes, que ya no deben someterse a unas normas dictadas por el operador de la plataforma central, y eso genera un mayor valor para el conjunto. Es importante destacar además que se trata de valor «nuevo», adicional al ya existente en la suma de cada uno de los participantes. Estas novedosas redes de usuarios, articuladas según ecosistemas de compartición y explotación de datos donde por primera vez se citan agentes que antes no parecían guardar una relación evidente, dan lugar a nuevos modelos de negocio con que generar valor adicional. Y esa dinámica debe hacerse de forma que no cree monopolios u oligopolios.

Por tanto, la gobernanza se convierte en un elemento esencial, debiendo establecerse reglas

para garantizar un funcionamiento adecuado, que sirvan para proteger los intereses de todos los participantes. En este sentido, hay que considerar y garantizar aspectos como la seguridad informática, la protección de la privacidad o la definición de los diferentes modelos de explotación y comercialización de datos. También como decimos, y de forma más rupturista frente al modelo centralizado, deben tomarse decisiones respecto al desarrollo y futuro del ecosistema, a la incorporación de nuevas funciones, la supresión de aquellas obsoletas y a la gestión de la comunidad de participantes. Esto es algo que en un modelo de plataforma no toman los usuarios, siendo el operador quien decide unilateralmente.

En el lado contrario, este modelo de ecosistema con gobernanza participativa lógicamente también supone nuevos retos y barreras en lo que respecta a la interoperabilidad: frente al paradigma de plataforma donde todos los participantes emplean el mismo vehículo (generalmente en forma de aplicación informática, y donde, por tanto, la interoperabilidad tecnológica y semántica existe desde el diseño), en el caso de un ecosistema federado deben primero acordarse las reglas técnicas con las que conectar entre sí los diferentes sistemas informáticos, que ahora pasan a estar distribuidos. Por ello, la capacidad para intercambiar y utilizar datos provenientes de diferentes sistemas pasa necesariamente por la homogeneización de modelos de datos (ya que estos suelen estar estructurados de manera diferente en diferentes sistemas), así como de los estándares de sistemas y plataformas que los gestionan, que pudieran no ser compatibles en origen. Para ello se pueden emplear diferentes marcos de referencia con especificaciones técnicas para el proceso del intercambio de datos, y el uso de APIs (interfaces de programación de ▷

aplicaciones) o servicios web que faciliten la integración entre diferentes sistemas. Asimismo, los vocabularios y ontologías compartidas ayudarán a estandarizar la estructura y semántica de los conjuntos de datos, facilitando su comprensión previamente al intercambio.

Pero, incluso más allá de la parte netamente tecnológica, deben además consensuarse unos códigos comunes de buenas prácticas que faciliten la interconexión entre organizaciones también a nivel de negocio, legal y organizativo, de forma que el intercambio de datos no genere cuellos de botella o impacte negativamente en los procesos de negocio que permite generarlos, compartirlos o explotarlos. Porque, de ser así, las organizaciones deberían acometer cambios para poder capitalizar estas nuevas lógicas comerciales en torno a los datos. Y con el alto número de combinaciones posibles entre las diferentes dinámicas estratégicas, estructurales y normativas bajo las que opera cada organización, la gran damnificada sería la liquidez del mercado único de datos, que se vería severamente limitado.

Es justo por este motivo que la Estrategia Europea de Datos busca una sólida coordinación entre las capacidades *cloud*, gracias a las que poder desplegar ágilmente servicios de infraestructura o *software* con los que adecuarse a las necesidades del momento, y el ámbito de la gestión, transformación y explotación de los datos. Sin dicha coherencia interactiva, los escenarios más probables serían uno donde los datos no lograsen romper sus silos originales, limitando por ende sus generosas capacidades transformativas; o, por el contrario, uno donde toda esa interoperabilidad por diseño venga dada por un grupo de interés con una posición dominante sobre el resto. En este sentido, el European Interoperability Framework (European Commission, Directorate-General for

Communications Networks, Content and Technology, and Directorate-General for Informatics, 2021) es un marco desarrollado por la Comisión Europea para promover la interoperabilidad en los sistemas de información utilizados por las Administraciones públicas. Si bien su alcance original es mucho más generalista, este resulta también una interesante opción para usarlo como referencia para el reto de la interoperabilidad en proyectos de compartición de datos, ya que proporciona principios, recomendaciones y buenas prácticas fácilmente aplicables.

En definitiva, el estudio de la interoperabilidad desempeña un papel crítico en el éxito de los ecosistemas de datos y en la aspiración de extraer todo el valor asociado a dichos entornos. Una de las principales peculiaridades de un sistema es la existencia de cualidades que resultan de la integración y contribución de cada uno de sus elementos, y cuyo valor excede a las características individuales, pero sin eliminarlas<sup>1</sup>. Y, para ello, la interoperabilidad es la pieza clave que articula dicha integración de elementos distribuidos, que debe por tanto crear unas condiciones que optimicen la dinámica natural de cualquier ecosistema de datos, favoreciendo la generación de sinergias entre sus participantes y donde estos maximicen sus cualidades y capacidades en tanto en cuanto conectan con el resto de los actores. La integración de capacidades (mediante nuevas conexiones) permite crear un valor «novedoso», derivado de las sinergias antes mencionadas, y explotar ese «valor latente en los datos» es precisamente lo que busca la Estrategia Europea de Datos mediante el despliegue de dichos ecosistemas federados. ▷

<sup>1</sup> Haciendo una referencia a la Teoría de Sistemas, «todo es más que la suma de las partes».



### 3. Los cuatro pilares de los ecosistemas federados de datos

Por definición, las plataformas digitales son susceptibles a los llamados efectos de red, aumentando rápidamente en valor al crecer su número de participantes. Esta bondad de las economías de escala resulta aún más beneficiosa al articularse alrededor de un esquema federado como el mencionado arriba, con un acoplamiento débil entre partes que busca incentivar y facilitar la participación, explotando los sistemas informáticos organizacionales sin costosas inversiones adicionales, o sin colocar a un intermediario en el centro que gobierne y determine el alcance del ecosistema. Sin embargo, y como comentábamos arriba, su modelo de gobernanza supone también un reto significativo. El mismo efecto que permite que cuanto más grande y activa sea una red, más valiosa se vuelve para todos sus miembros (creando un efecto multiplicador en términos de oportunidades de colaboración, conocimiento compartido y valor generado con base en sinergias), implica también una necesidad más acuciante a la hora de coordinar dicha interconexión entre múltiples actores de forma adecuada.

Por este motivo, postulamos cuatro líneas estratégicas sobre las que fomentar la creación y operativa de estos ecosistemas soberanos, con el objetivo de dinamizar su acogida y ganar capilaridad.

1. Dado el carácter distribuido de estos ecosistemas, la generación de comunidad resulta un elemento fundamental. En cualquier sector existen elementos comunes y diferenciales únicos al mismo, resultado de sus procesos de negocio específicos, así como de una progresiva

e iterativa generación de conocimiento y una cultura común. Es por ello por lo que cualquier actuación dirigida al impulso de los ecosistemas federados de datos debe considerar no sólo las particularidades de la comunidad sobre la que busca encontrar despliegue, sino también desarrollarse sobre los marcos preexistentes en la misma, reutilizando artefactos, estándares, mejores prácticas y códigos de conducta de amplia aceptación en dicho ámbito. Dicha comunidad no solamente debe ser tenida en cuenta para llegar a buen puerto, sino también potenciada.

Para este fin sugerimos varias acciones: (i) el fomento de la comunicación y la interacción entre participantes, proporcionando canales de intercambio de información y plataformas colaborativas donde los miembros de un ecosistema puedan compartir conocimiento, experiencias y prácticas; (ii) el establecimiento de refuerzos a la participación, por ejemplo, mediante el uso de incentivos económicos, acceso privilegiado a recursos y servicios, o reconocimientos públicos; (iii) la creación de alianzas estratégicas, articuladas bien alrededor del establecimiento de acuerdos comerciales o con base en la creación de proyectos conjuntos; (iv) proporcionar apoyo y recursos, incluyendo herramientas y capacitación que ayuden a los miembros de un ecosistema a desarrollar proyectos, asesoramiento técnico específico, orientación sobre posibles modelos de negocio o la «disponibilidad» de infraestructuras tecnológicas compartidas — dedicaremos más adelante un apartado especial a este tema—; y (v) ▷

establecer procesos de gobernanza participativa que favorezcan que los miembros de la comunidad se involucren en la toma de decisiones y en la evolución del ecosistema, por ejemplo, a través de mecanismos de retroalimentación como encuestas, grupos de trabajo y, por supuesto, también mediante la participación de representantes de la comunidad en los órganos de gobernanza del mismo.

Creemos que todas estas clases de iniciativas ayudarán firmemente en la fase de adopción inicial, durante la que el valor de la red puede no ser evidente o significativo. Por ello, se hace necesario atacar ámbitos de actuación conjuntos y colaborativos con los que fomentar la participación en un ecosistema federado de datos, superando barreras iniciales y generando el impulso necesario para el crecimiento y el éxito de la red. Creemos que una vez que las organizaciones comiencen a unirse, crecerá también el incentivo para que otras les sigan.

2. Como comentábamos arriba, la interconexión entre participantes es el elemento central de cualquier red, también en una de datos. Y dado que estas tienen como objetivo alcanzar objetivos comunes y generar un valor conjunto, se basan en relaciones de confianza y cooperación mutua. Y es precisamente alrededor del reto de crear marcos de confianza donde se encuentra uno de los elementos nucleares de los proyectos de compartición de datos.

Y dado que en los modelos de ecosistemas la centralización propia de las plataformas se rompe, la transparencia y la

confianza adquieren en estos un rol incluso aún más preponderante, al no poder reunir a todos los participantes de forma simultánea, debido a su naturaleza distribuida. Por tanto, cualquier iniciativa a desplegar debe considerar las políticas y mecanismos que promulguen adecuadamente estas dimensiones. En este sentido, creemos que diversas actuaciones pueden ayudar a tal fin: (i) el diseño y consenso de un marco común operativo, que también incluya aspectos legales y normativos; (ii) la comunicación y actualización a los participantes sobre cambios y decisiones tomadas en el ecosistema; (iii) el despliegue de herramientas técnicas que permitan una correcta identificación y acreditación de participantes y servicios ofertados; (iv) el uso de técnicas informáticas avanzadas para la monitorización y control del uso de los datos, una vez estos han sido compartidos; y (v) la definición y el empleo de protocolos tecnológicos que favorezcan la automatización de las interacciones.

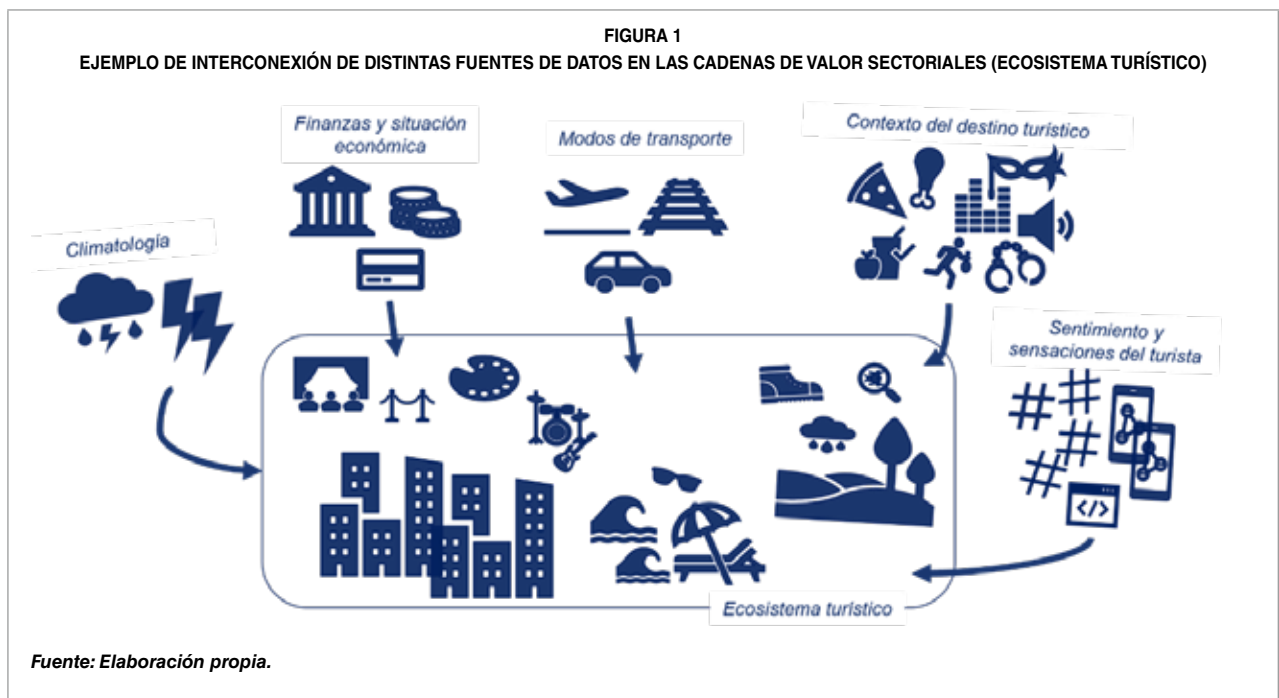
Todas estas medidas fortalecen y cohesionan los ecosistemas, y buscan resolver una parte importante de los retos alrededor del principio de federación que introduce el modelo de ecosistemas de datos a la UE. Los seres humanos, y las organizaciones como extensión, operamos sobre la base de la confianza, y para que cualquier diseño de mercado adquiera tracción y gane capilaridad, esta debe ser celebrada y promovida.

3. La modelización *data-driven*, también a menudo llamada ciencia de datos, es ▷



la disciplina del análisis de datos para la toma de decisiones y para la mejora de los procesos de negocio. Lógicamente, en este campo los mismos constituyen el elemento basal de todo el proceso, utilizándose para identificar patrones y tendencias que permitan mejorar la efectividad y la eficiencia en las organizaciones. Esto ha llevado a diversos analistas a referirse al dato como «el petróleo del siglo XXI», dada su enorme plasticidad y capacidad transformativa a lo largo de diferentes contextos, y por el rol primordial que juega en la innovación de productos y servicios. Dicha capacidad disruptiva es la seña de identidad de las organizaciones modernas impulsadas por los datos, que orientan sus procesos (incluidos los más tradicionales o preexistentes) en torno al dato, como base para mejorar sus modelos operativos.

Esta disciplina, si bien todavía pendiente de aplicación a gran escala, no es necesariamente nueva. Se lleva usando en entornos acotados de tipo médico, epidemiológico, industrial o estadístico durante más de cien años, y lógicamente recibió un fuerte impulso en las últimas décadas con la llegada de la computación programática y la «comoditización» de los sistemas informáticos. Más recientemente, ya en el siglo actual, el desarrollo de innovadoras arquitecturas tecnológicas para un procesamiento en paralelo (asequible) ha permitido la democratización del campo, haciendo con ello que los expertos técnicos sean altamente demandados. Estamos, por tanto, en una era caracterizada por la gigantesca capacidad de generar datos, y con infraestructuras digitales que en teoría pueden dar abasto a tales volúmenes (al menos si el problema a resolver se acota razonablemente). ▷



Sin embargo, el proceso de interconectar dichas fuentes de datos entre sí, a escala, es algo que todavía sigue sin resolverse adecuadamente. Y es precisamente para ello que surgen los ecosistemas federados, buscando romper los silos naturales en los que los datos han sido recogidos y explotados de forma primaria. Esto es tremendamente innovador, ya que, al considerar los datos como un insumo productivo, existen numerosas acepciones bajo las que un mismo conjunto puede ser analizado al interrelacionarlo con otros del mismo ámbito, y sobre todo con conjuntos de ámbitos complementarios, en cuyo caso las oportunidades crecen exponencialmente. Esta proposición de valor es la perspectiva aspiracional desde la que abordar el desarrollo de los ecosistemas federados de datos, con base en la alta escalabilidad y flexibilidad del arquetipo propuesto.

Con este objetivo en mente, sugerimos varias acciones concretas para impulsar la innovación en base a datos, como (i) el empleo de metodologías de ideación, con que teorizar soluciones según la compartición de conjuntos de datos; (ii) la capacidad de pilotar y desplegar ágilmente pruebas de concepto, con un impacto instrumental en el desarrollo de estos ecosistemas federados donde —por diseño— no existe una figura predominante intermediando el sistema a la que acogerse; y (iii) la elaboración de estudios sobre la madurez en torno a la captura y explotación del dato a lo largo de diferentes sectores productivos, así como un análisis prospectivo de actuaciones recomendadas según los resultados.

4. Por último, y para que un ecosistema de datos goce de practicidad y continuidad, este debe evolucionar de la fase piloto a una realidad operativa desde donde generar un valor de negocio cuantificable, es decir, alcanzar lo que en *lingo* inversionista se suele llamar el *scale-up*. Y esto pasa indudablemente por el despliegue de procesos sostenibles, desarrollados sobre dinámicas y escenarios de negocio y partenariados que den pie a un beneficio compartido, y cuyas bases organizativas, legales y regulatorias gocen de coherencia y solidez.

Las dinámicas de construcción de una propuesta de valor (claramente percibida y apreciada por los miembros del ecosistema) son clave para la sostenibilidad de estos. Dado que los ecosistemas tienden a generar un mayor valor que la mera suma de sus partes, resulta fundamental crear las condiciones que garanticen su continuidad en el tiempo, pues de otro modo las sinergias entre participantes desaparecerán, y, con ellas, la razón misma de la existencia del ecosistema. En este sentido, sugerimos algunas iniciativas para alcanzar un impacto positivo a medio y largo plazo: (i) el empleo de mecanismos de monitorización y trazabilidad de procesos, incluyendo funciones de «cámara de compensación y auditoría»; (ii) la identificación y el consenso en torno a un modelo de gobernanza para la resolución de incidentes y disputas, y el despliegue de una unidad operativa que lo implemente; (iii) el empleo de normas y estándares certificables en materia de gestión y calidad del dato (e.g., tipo UNE o ISO) con los que eliminar barreras ▷

para los participantes del ecosistema e impulsar un insumo productivo de mayor calidad; o (iv) la promoción de modelos cooperativos en torno a datos, que ayuden a aquellas organizaciones con una menor madurez o una huella de negocio y talento más limitada, como las pymes.

#### 4. Infraestructura común compartida

Estas cuatro líneas estratégicas, ámbitos sobre los que desplegar los ecosistemas federados de datos, son dominios que ya han sido analizados extensamente de forma individual con anterioridad. Por tanto, la novedad en este caso surge de conjugarlos mutuamente entre sí. Por ejemplo, en entornos centralizados o de plataforma, la conversación suele girar en torno a la innovación con base en datos y sus consideraciones de cara a escalar dichos modelos *data-driven* (dando por sentado la disponibilidad y eficacia de las infraestructuras subyacentes y que los participantes mantienen una relación de confianza con el operador). Sin embargo, en el contexto de los entornos federados, estas dos líneas (más cercanas al ámbito de los datos) deben además considerar la capacidad de interconectar fuentes de datos localizadas en sistemas de información de distintos orígenes y dueños, y cuya reutilización genera, por tanto, disyuntivas respecto a su origen y semántica, si son siquiera confiables y en lo relativo a bajo qué condiciones mercantiles, legales u organizativas rentarían.

Por este motivo, la orquestación transparente de participantes y recursos de datos se alza como un elemento central en la creación de los ecosistemas federados. Idealmente, esta coordinación se aplicará a lo largo de tantos procesos transformativos y de explotación del dato

como sea posible, puesto que, de existir puntos de fallo o cuellos de botella en la interrelación entre participantes, esto penalizará al modelo de negocio e irá en detrimento de la percepción del propio ecosistema. En este sentido, también creemos que se debe preponderar el uso de las herramientas y procesos ya desplegados en la arquitectura informacional de las diferentes organizaciones participantes, dado que esto facilita la cohesión organizativa y elimina fricciones de entrada. Por lo tanto, y aunque el modelo de plataforma sí habilita una adecuada orquestación, gracias a una gobernanza e infraestructura común a todos los participantes, sufre en lo que respecta al segundo punto, pues dicha infraestructura solo le pertenece al operador, a la que deben adaptarse las organizaciones que hacen uso de ella, y sobre la que puede tomar decisiones unilateralmente.

Por el contrario, el modelo que proponemos nosotros hace uso de una infraestructura común que no está integrada monolíticamente. Es decir, en lugar de abarcar todos los componentes técnicos, la gobernanza y las aplicaciones con que usar los datos, solamente prescribe la capacidad de interconexión de recursos y procesos, siempre bajo la premisa de un enfoque federado y auspiciando el respeto a la autodeterminación de los agentes intervinientes. Es decir, se busca potenciar las capacidades y cualidades ya existentes en los participantes del propio ecosistema, y maximizarlas con base en ingredientes de interoperabilidad que permitan que las dinámicas naturales de interacción se realicen de forma más extensiva y dinámicamente. Eso generalmente implica:

- Incorporar el mayor número de actores al ecosistema, buscando también que se fomente la integración del mayor número posible de capacidades. ▷

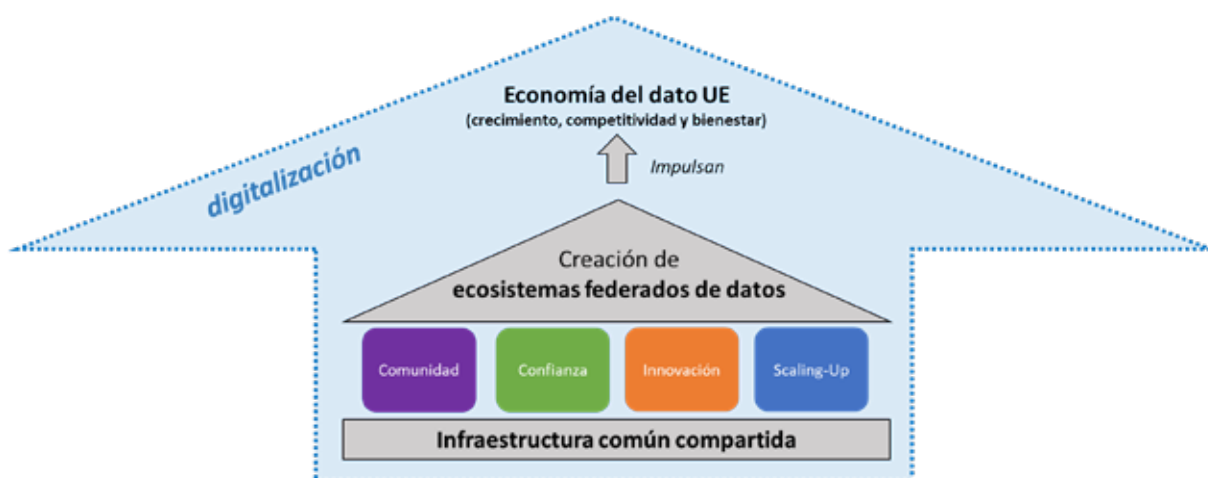
- Asegurar que el ecosistema sea representativo de la cadena de valor sobre la que se quieren aprovechar las oportunidades de la nueva Economía del Dato.
- Reutilizar las capacidades existentes para construir valor agregado (y no destruir valor previamente construido), ayudando a los participantes a amortizar sus inversiones previas en conocimiento, tecnología y talento.

Así, análogamente a como internet surgió para ganar resiliencia operativa (en las comunicaciones entre ciudades) gracias a un modelo distribuido, la creación de una capa de «infraestructura común compartida» (Dion y Pons, 2022) permite desplegar la deseada orquestación transparente, confiable y eficiente entre diferentes combinaciones de los muchos y potenciales participantes de los ecosistemas federados de datos. Podríamos decir además que se potencia una descentralización del poder, democratizando el acceso a las oportunidades de la Economía del Dato, y nivela el campo de

juego para cualquier organización, con independencia de su tamaño y capacidades. Y no lo hace de forma verticalizada alrededor de un dominio concreto, como puede ser el caso de ofertas de plataformas sectoriales monolíticas ya disponibles, sino con base en una capacidad de interconexión descentralizada entre la oferta y la demanda de servicios y recursos de diferentes proveedores, colectivizando el valor entre diferentes grupos de interés de características heterogéneas, que han pasado así a estar acoplados de forma soberana y débil.

Este esquema, asimilable a las capacidades transversales de una red de infraestructuras fundamentales de un territorio (luz, agua, saneamiento, carreteras, etc.), busca generar las condiciones favorables para el desarrollo de la capa más básica del ansiado mercado único de datos a escala europea, aportando herramientas compartidas con que suscitar economías de red, reducir barreras de entrada de participantes e impulsar las capacidades de innovación y resiliencia de las industrias de la Unión. Sin embargo, esta no tiene ▷

FIGURA 2  
INFRAESTRUCTURA COMÚN COMPARTIDA COMO HABILITADORA DE LAS OPORTUNIDADES DE LA ECONOMÍA DEL DATO



Fuente: Elaboración propia.

solamente una representación física (lo que se suele llamar *hard infrastructure*), como es, por ejemplo, el desarrollo de laboratorios, entornos de desarrollo de pruebas, aplicaciones específicas de transformación o consumo de datos, o infraestructuras de ejecución (*run-time environments*), sino que la infraestructura común adopta también caracterizaciones más suaves. En esta *soft infrastructure* se incluye el desarrollo de estándares y mecanismos de conformidad, piezas *software* comunes reutilizables<sup>2</sup> o pruebas piloto a lo largo de diferentes dominios. También pueden considerarse intangibles como la coordinación de los ecosistemas y la dinamización e incubación de las comunidades y sus participantes, así como el impulso a la reutilización de los datos abiertos en posesión de las Administraciones públicas.

Para que verdaderamente ofrezca respuestas en materia de interoperabilidad, toda esta infraestructura común debe buscar acomodo a lo largo de la «dimensión de negocio», con base en el análisis de modelos económicos y organizativos, y el impulso a la cooperación y la innovación colaborativa, y de la «dimensión legal», ofreciendo respuestas a las consideraciones y necesidades contractuales y regulatorias de los participantes. También debe considerar la funcional y operacional, incluyendo catálogos de recursos disponibles bajo este esquema federado, el impulso y la dinamización de participantes (a fin de generar liquidez y un amplio abanico de servicios), la caracterización de roles y mejores prácticas a ejercer o la capacitación y el despliegue de comunidades de soporte que atesoren y avancen un conocimiento común compartido.

En resumen, si bien la mutualización de estos desarrollos y artefactos no es la única pieza

necesaria para la creación de los ecosistemas federados de datos, sirven para dar forma a una infraestructura horizontal sólida y sostenible, común a nivel España y Unión Europea, y en línea con los denominados **Bienes Públicos Digitales**. Creemos que esta infraestructura servirá como piedra angular de la incipiente y ambiciosa Economía del Dato europea, evitando la duplicación de esfuerzos y fomentando la colaboración con base en un esfuerzo y un retorno distribuido, transversal e inclusivo. Y respaldando una visión humanista que sitúa a la ciudadanía en el centro de la política digital.

## Bibliografía

Dion, O., & Pons, A. (2022). *Data de Confiance. Le partage des données, clé de notre autonomie stratégique*. Digital New Deal. <https://www.thedigitalnewdeal.org/data-de-confiance-le-partage-des-donnees-cle-de-notre-autonomie-strategique/>

European Commission, Directorate-General for Communications Networks, Content and Technology, Cattaneo, G., Micheletti, G., Glennon, M., La Croce, C., & Mitta, C. (2020). *The European data market monitoring tool. Key facts & figures, first policy conclusions, data landscape and quantified stories. D2.9 Final Study Report*. Publications Office of the European Union. <https://data.europa.eu/doi/10.2759/72084>

European Commission, Directorate-General for Communications Networks, Content and Technology, and Directorate-General for Informatics. (2021). *Proposal for a European Interoperability Framework for Smart Cities and Communities (EIF4SCC)*. Publications Office of the European Union. <https://data.europa.eu/doi/10.2799/816559>

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet Classification with Deep Convolutional Neural Networks. *Advances in Neural* ▷

<sup>2</sup> Disponibles modularmente en forma de código de fuentes abiertas, y que generalmente van a requerir de trabajo de adaptación o integración adicional, previo a su uso.

Alberto Palomo Lozano y Manuel Gutiérrez de Diego

*Information Processing Systems*, 25. [https://proceedings.neurips.cc/paper\\_files/paper/2012/file/c399862d3b9d6b76c8436e924a-68c45b-Paper.pdf](https://proceedings.neurips.cc/paper_files/paper/2012/file/c399862d3b9d6b76c8436e924a-68c45b-Paper.pdf)

### **Página web**

The Digital Public Goods Alliance. <https://digitalpublic-goods.net/>