



Radiografía de Tecnologías en Proyectos de Espacios de Datos.



V1 - 15 de octubre de 2025



Financiado por
la Unión Europea
NextGenerationEU



GOBIERNO
DE ESPAÑA
MINISTERIO
PARA LA TRANSFORMACIÓN DIGITAL
Y DE LA FUNCIÓN PÚBLICA

SECRETARÍA DE ESTADO
DE DIGITALIZACIÓN
E INTELIGENCIA ARTIFICIAL



Plan de Recuperación,
Transformación
y Resiliencia

Contenido

1. Descripción general.....	3
2. Metodología	3
3. Análisis Técnico	5
3.1. Fase actual del proyecto.....	5
3.2. Nivel de criticidad de los datos	6
3.3. Infraestructura utilizada y desglose de Cloud	7
3.4. Principales proveedores tecnológicos (integradores).....	8
3.5. Proveedores tecnológicos de software o datos.....	9
3.6. Arquitecturas de referencia o Building Blocks utilizados	10
3.7. Modelos y tecnologías que utilizan los datos que se suministran	11
3.8. Técnicas de comunicación que se utilizan para la transferencia de datos ...	11
3.9. Soluciones y/o servicios que se utilizan para la transferencia de datos	12
3.10. Gestión de confianza en el espacio de datos.....	13
3.11. Soluciones y/o servicios que se utilizan para el control de la ejecución de las políticas de uso de datos.....	14
3.12. Soluciones y tecnologías que se utilizan para garantizar la seguridad y privacidad.....	15
3.13. Herramientas y/o servicios que se utilizan para la catalogación de productos de datos, servicios y otros recursos.....	16
5. Anexo A. Cuestionario	18

1. Descripción general

La implementación de espacios de datos representa un desafío tecnológico complejo que requiere la orquestación de múltiples componentes: desde la gestión de identidades y el control de acceso, hasta la transferencia segura de datos y el cumplimiento de políticas de uso.

El Centro de Referencia de Espacios de Datos, y enmarcado en su iniciativa RADAR, ha elaborado el estudio "**Radiografía de tecnologías en Espacios de datos**", que ha sido diseñado con los siguientes objetivos:

1. Analizar el ecosistema tecnológico actual de los proyectos de espacios de datos beneficiarios de ayudas públicas, identificando patrones, buenas prácticas y áreas de mejora.
2. Identificar oportunidades de estandarización y reutilización de soluciones.
3. Detectar carencias tecnológicas (gaps) y necesidades de capacitación.
4. Facilitar la interoperabilidad entre espacios de datos mediante el uso de estándares comunes.

El análisis de los resultados permite obtener una visión estructurada del estado del arte tecnológico, detectar tendencias emergentes y proporcionar recomendaciones técnicas orientadas a la optimización y convergencia de futuras implementaciones. Permite establecer una base empírica para orientar decisiones estratégicas, definir líneas de apoyo técnico y promover la convergencia tecnológica entre actores del ecosistema.

El análisis de los resultados se centra en dimensiones clave del diseño e implementación tecnológica de espacios de datos, incluyendo gestión de identidades y control de acceso, aplicación de políticas de uso de datos, seguridad y privacidad, catalogación de datos y recursos, estándares e interoperabilidad, reutilización y sostenibilidad tecnológica.

Este informe sintetiza los hallazgos obtenidos a partir de las respuestas recopiladas y ofrece una visión estructurada y comparativa del panorama actual de los espacios de datos en desarrollo o en operación en España.

2. Metodología

Para la elaboración de este informe se diseñó y distribuyó un cuestionario estructurado dirigido a los proyectos beneficiarios de iniciativas relacionadas con espacios de datos de la primera convocatoria de demostradores y casos de uso de espacios de datos realizada por la Secretaría de Estado de Digitalización e Inteligencia Artificial.

El objetivo fue recopilar información detallada sobre las tecnologías, arquitecturas, proveedores y prácticas implementadas en cada proyecto.

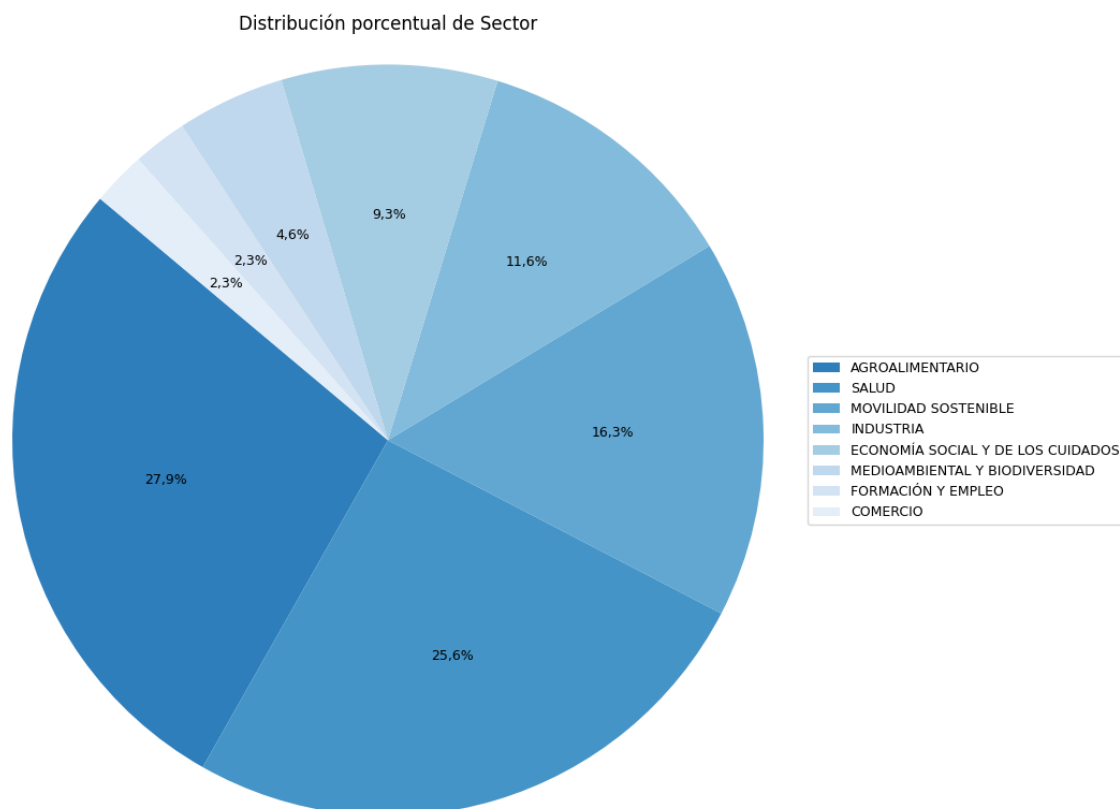
El cuestionario constó de 15 bloques temáticos, con preguntas cerradas y de opción múltiple, abarcando aspectos como:

- Fase de desarrollo del proyecto.
- Nivel de criticidad de los datos tratados.
- Infraestructura tecnológica utilizada.
- Proveedores tecnológicos e integradores.
- Arquitecturas de referencia adoptadas.
- Modelos de datos y técnicas de comunicación.
- Soluciones de transferencia de datos y conectores.
- Mecanismos de gestión de la confianza.
- Control de políticas de uso de datos.
- Tecnologías de seguridad y privacidad.

- Herramientas de catalogación de datos y servicios.

Un total de **50 entidades** han enviado contribuciones a este cuestionario. De ellas, la mayoría —**27 de 50**— corresponden a **empresas privadas**. El resto incluye **23 entidades de otro tipo**, entre las que se encuentran **organismos públicos de investigación (18)**, **asociaciones empresariales (4)** y **organismos privados de investigación (1)**. Los datos fueron procesados y analizados mediante técnicas de análisis descriptivo, identificando frecuencias, patrones de uso y combinaciones tecnológicas. Las respuestas abiertas fueron categorizadas para facilitar su interpretación.

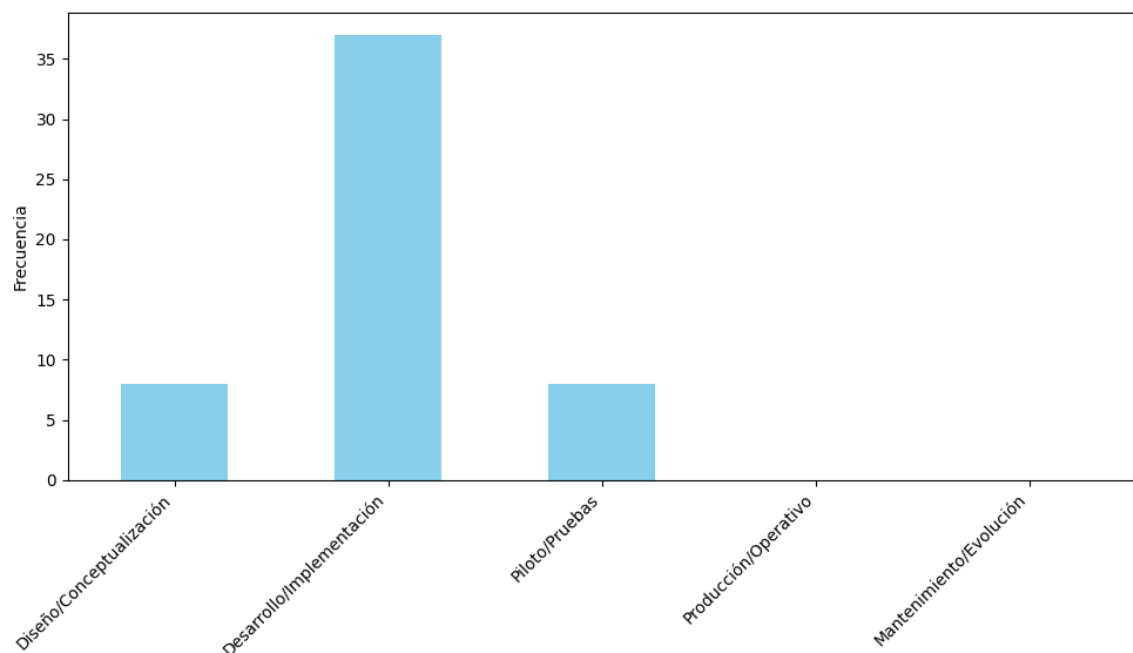
Por sector, estas entidades se encuadran, en su mayoría, en el sector Agroalimentario, con un 28,3%; el de Salud, con un 17,4%; Movilidad sostenible, con un 15,2%; y en Industria con un 15%.



Se observa una concentración significativa en dos provincias, Madrid y Barcelona, que en conjunto agrupan el 34,69% del total. El 65,31% restante se reparte entre 16 provincias, lo que refleja una participación territorial diversa. Si bien esta distribución indica cierto grado de descentralización, también evidencia la necesidad de continuar impulsando la implantación y madurez de proyectos en otras regiones para fomentar un desarrollo más equilibrado.

Por otro lado, un 15 % de los proyectos está en **fase de piloto o pruebas**, indicando que se encuentran en procesos de validación y ajuste antes de escalar o desplegarse en producción. De forma complementaria, otro 15 % permanece en la **fase de diseño o conceptualización**, evidenciando la presencia de iniciativas más incipientes en proceso de definición.

En conjunto, esta distribución refleja un **ecosistema dinámico y en evolución**, donde conviven proyectos en distintas fases del ciclo de vida. Esta diversidad permite una renovación continua del conjunto de iniciativas en marcha, combinando iniciativas consolidadas con otras emergentes que alimentan futuras etapas de desarrollo.



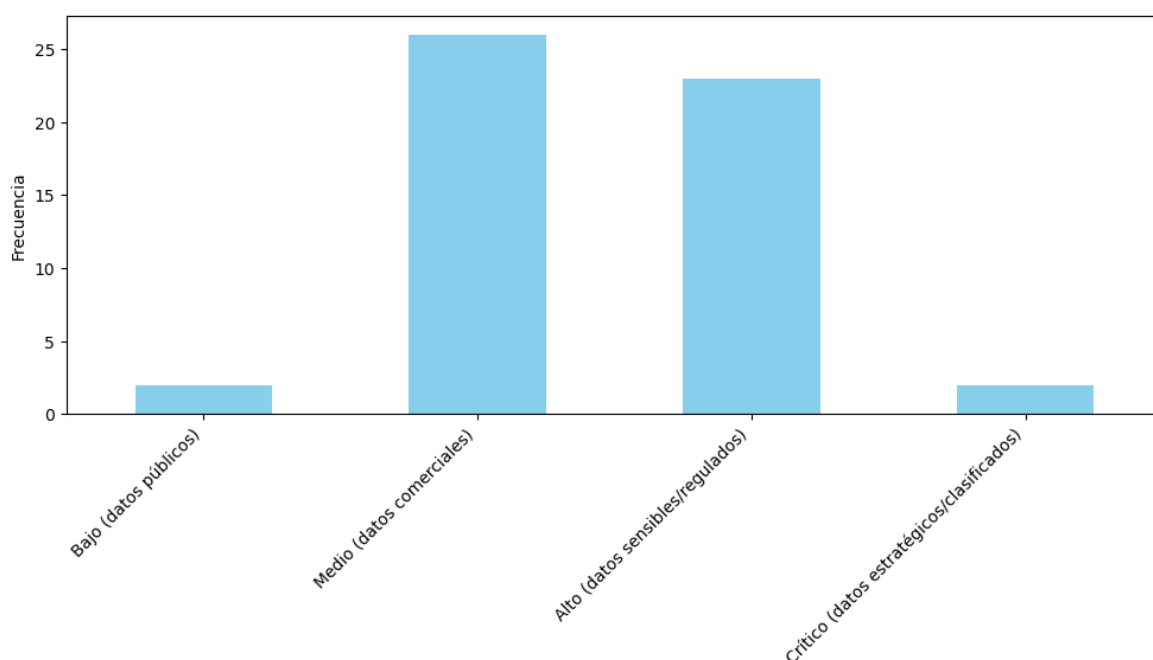
3.2. Nivel de criticidad de los datos

Los resultados reflejan que la gran mayoría de los proyectos opera con **datos de criticidad media o alta**, lo que sitúa la **seguridad de la información y el cumplimiento normativo** como elementos clave en el diseño y operación de estos espacios de datos.

Casi la mitad de las iniciativas, 49 %, trabaja con **datos comerciales**, lo que implica la necesidad de proteger activos de valor económico y garantizar acuerdos de uso definidos. A su vez, un 43 % gestiona **datos sensibles o sujetos a regulación**, como los de carácter personal o sectorialmente restringidos, lo que exige medidas robustas de **trazabilidad, cifrado, control de acceso y auditoría**.

En contraste, solo un **7,5 %** de los proyectos opera con **datos públicos o estratégicos**, lo que sugiere que la mayoría de los espacios de datos se desarrollan en **contextos cerrados o regulados**, donde **la confidencialidad, la integridad y la disponibilidad** de la información son prioridades operativas.

Esta distribución pone de relieve la importancia de contar con **infraestructuras seguras y marcos de gobernanza sólidos**, especialmente en entornos colaborativos donde intervienen múltiples actores y se comparten datos de alto valor.



3.3. Infraestructura utilizada y desglose de Cloud

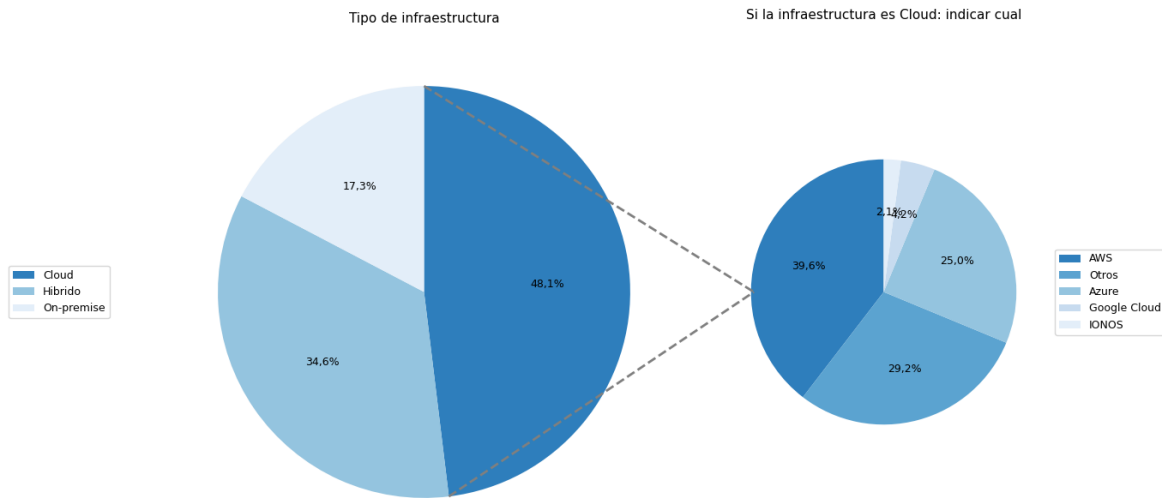
Los resultados muestran que casi la mitad de los proyectos, 48 %, utiliza infraestructura en la nube, lo que refleja una clara apuesta por entornos **escalables, flexibles y de rápida implementación**, especialmente relevantes en contextos donde la interoperabilidad y la colaboración entre entidades son clave.

Un 35 % adicional se apoya en arquitecturas **híbridas**, combinando servicios Cloud con recursos locales (on-premise), lo que responde a necesidades de **compatibilidad con sistemas existentes** o **control sobre información sensible**. El 17 % restante opera exclusivamente en entornos locales, lo que puede estar motivado por **exigencias regulatorias, de seguridad o políticas internas**.

Dentro del grupo de proyectos que utilizan servicios en la nube (ya sea de forma total o parcial), **AWS** es el proveedor más empleado, seguido de **Microsoft Azure**.

La presencia significativa de la categoría **“Otros”** indica que algunos proyectos están optando por proveedores más pequeños o especializados, posiblemente motivados por **criterios de cumplimiento normativo local, sostenibilidad, costes, etc.**

Infraestructura utilizada y desglose de Cloud



3.4. Principales proveedores tecnológicos (integradores)

Una proporción significativa de los proyectos, casi el **40 %**, declara utilizar **medios propios** para el desarrollo e integración de sus soluciones. Esta tendencia puede interpretarse como una apuesta por la **autonomía tecnológica**, el aprovechamiento de **capacidades internas ya disponibles** o bien como una señal de que algunos proyectos se encuentran todavía en **fases tempranas**, en las que no se ha formalizado la contratación de integradores externos.

El resto de las respuestas revela un panorama **altamente fragmentado**, con una amplia variedad de proveedores mencionados, la mayoría de ellos citados solo una vez. Únicamente cuatro entidades —**Telefónica**, **Gantabi**, **Purple Blob** y **Tecnalia**— aparecen en más de un proyecto.

Esta dispersión sugiere que las decisiones sobre proveedores se están tomando de forma **muy contextual**, en función de factores como la **localización**, el **sector de aplicación**, el **estado del proyecto** o la **experiencia previa** con determinados socios tecnológicos.

El panorama actual refleja un ecosistema en construcción, donde los proyectos tienden a personalizar sus alianzas tecnológicas y aún no hay una consolidación clara de proveedores líderes. A medida que el sector madure, es probable que algunos integradores adquieran mayor relevancia y aparezcan con mayor frecuencia como socios tecnológicos preferentes.

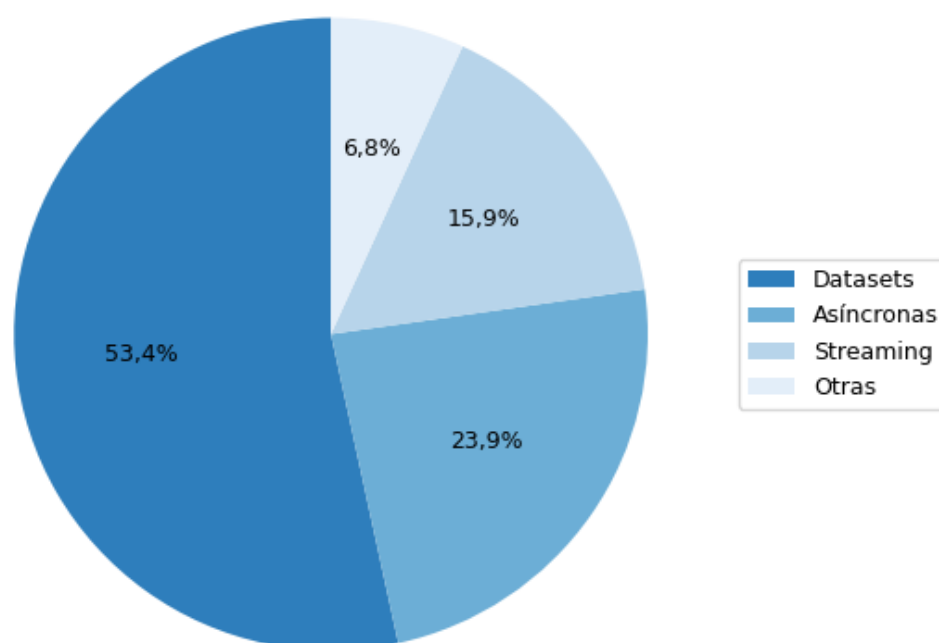
3.7. Modelos y tecnologías que utilizan los datos que se suministran

Más de la mitad de los proyectos, 53,4%, utilizan **datasets** como modelo principal para suministrar datos, lo que indica una clara preferencia por estructuras de datos estáticas o procesadas por lotes. Este enfoque es habitual para análisis tradicionales, interoperabilidad y publicación de datos.

Además, un porcentaje significativo de proyectos emplea modelos **asíncronos**, 23,9%, y **streaming**, 15,9%. Estos enfoques revelan una creciente adopción de arquitecturas orientadas a eventos y procesamiento de datos en tiempo real, vitales en sectores que demandan inmediatez como movilidad, salud o industria.

Por último, el 6,8% restante usa otros modelos, mostrando que existe cierta exploración hacia tecnologías y métodos específicos o alternativos según las necesidades de cada sector.

Los datos que se suministran, ¿Qué modelos y tecnologías utilizan?



3.8. Técnicas de comunicación que se utilizan para la transferencia de datos

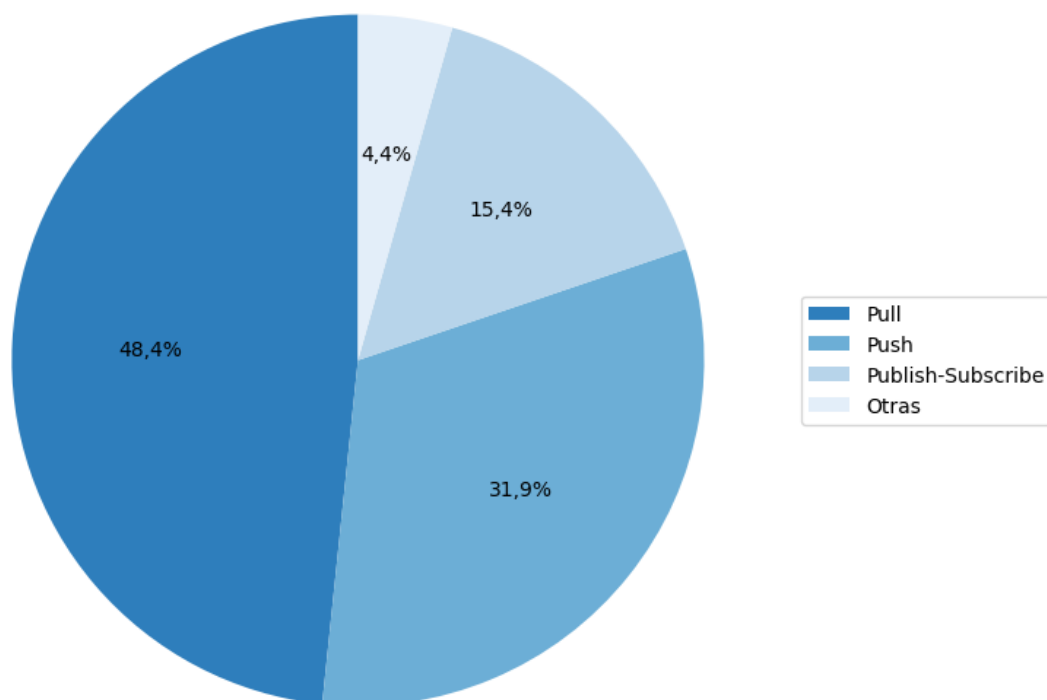
La técnica **Pull** es la más utilizada, con una presencia en el 48,4% de los proyectos. Este modelo, donde los consumidores solicitan los datos activamente, favorece la trazabilidad, el control de acceso y la eficiencia en entornos donde los datos no cambian constantemente o donde se busca minimizar el tráfico innecesario.

La técnica **Push**, presente en el 31,9% de los proyectos, implica el envío automático de datos a los consumidores, lo que es útil en contextos que requieren actualizaciones frecuentes o en tiempo real, como sensores, IoT o flujos continuos de información.

El modelo **Publicar y Suscribirse**, 15,38%, muestra una adopción más moderada, pero relevante en escenarios donde se requiere una arquitectura distribuida y reactiva, como sistemas de notificación o intercambio de eventos entre múltiples actores.

Finalmente, un 4,4% de los proyectos emplea **otras técnicas**, lo que puede incluir enfoques personalizados o específicos del sector, como APIs multipunto, soluciones propietarias o infraestructuras especializadas (e.g., redes VAN).

¿Qué técnicas de comunicación se utilizan para la transferencia de datos?



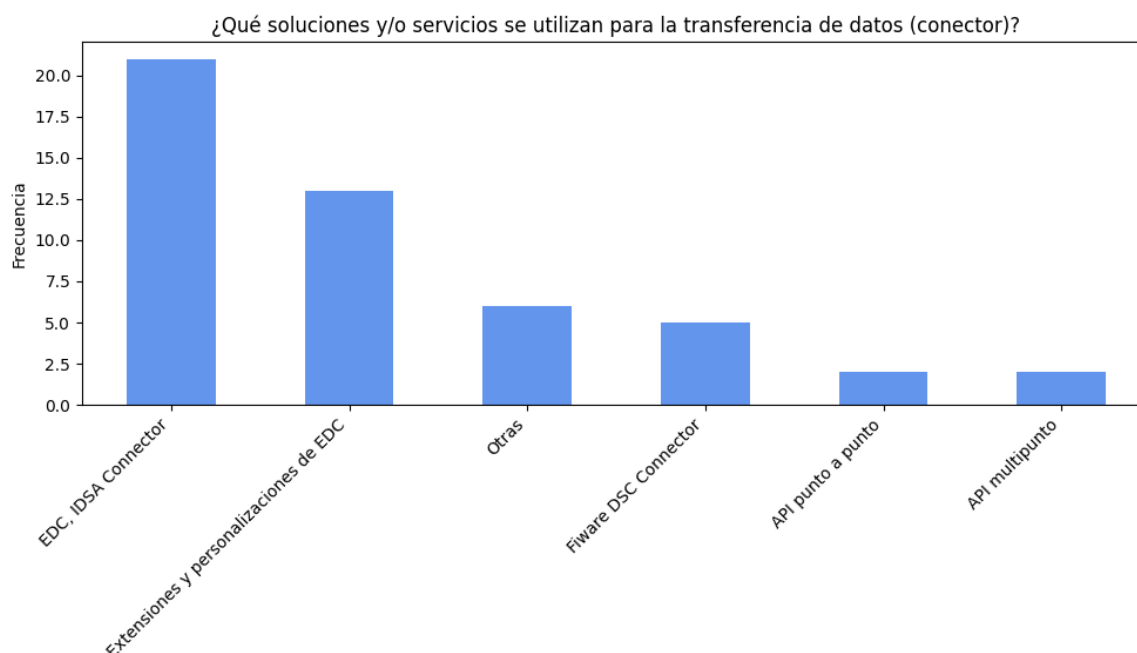
3.9. Soluciones y/o servicios que se utilizan para la transferencia de datos

El conector “**EDC/IDSA**” es mencionado en el **42,86%** de los proyectos como solución utilizada para la transferencia de datos. Un **26,5%** de los casos indica el uso de “extensiones o personalizaciones” de este conector.

El “**FIWARE DSC Connector**” aparece en el **10,2%** de las respuestas. Las “**APIs punto a punto**” y “**multipunto**” son mencionadas en un **4,08%** de los casos cada una.

Finalmente, la categoría “**Otras**”, 12,24%, refleja la diversidad tecnológica que aún existe en el ecosistema, donde algunos proyectos están explorando soluciones propias

o alternativas, posiblemente ligadas a especificidades sectoriales o a etapas tempranas de desarrollo.



3.10. Gestión de confianza en el espacio de datos

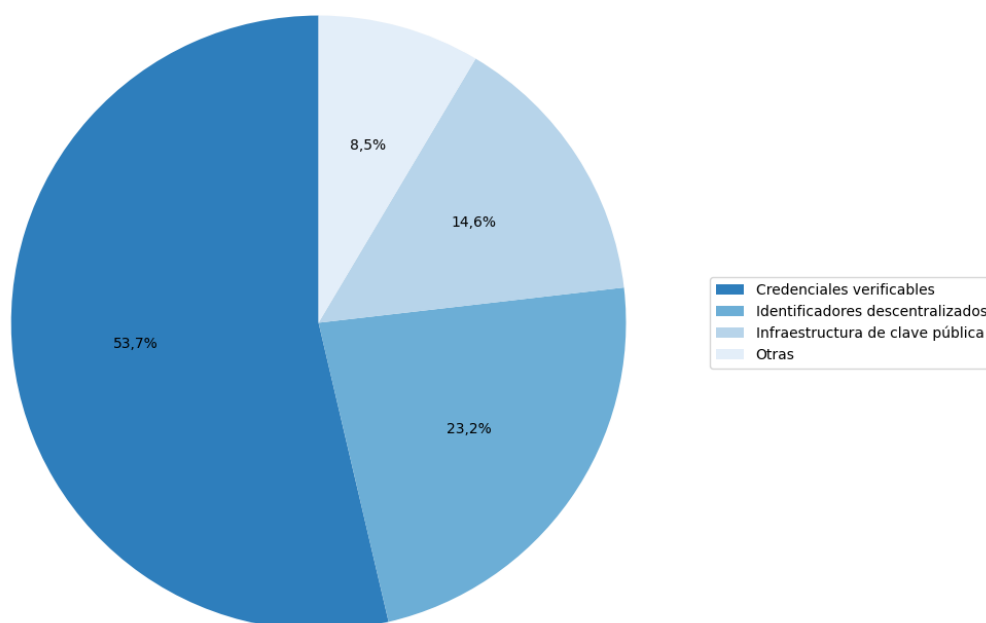
La **gestión de la confianza** en los espacios de datos se basa principalmente en el uso de **credenciales verificables**, adoptadas por más de la mitad de los proyectos, 53,66%. Esta prevalencia evidencia una clara orientación hacia modelos que permiten garantizar de forma segura la autenticidad y la autorización de los actores involucrados, alineándose con los principios de identidad digital y trazabilidad definidos en los marcos europeos.

Los **identificadores descentralizados (DIDs)**, utilizados en el 23,17% de los casos, reflejan una adopción significativa de enfoques basados en **identidad soberana**. Estos modelos son especialmente relevantes en arquitecturas federadas, donde no existe una autoridad central y se requiere asegurar la autonomía y privacidad de los participantes.

La **infraestructura de clave pública (PKI)** aparece como solución en el 14,63% de los proyectos, lo que sugiere que muchos entornos complementan mecanismos más avanzados (como credenciales verificables o DIDs) con tecnologías criptográficas tradicionales que siguen siendo fundamentales para el cifrado, la autenticación y la firma digital.

Por último, un 8,54% menciona el uso de **otras soluciones**, posiblemente relacionadas con marcos propios de confianza, tecnologías emergentes o estrategias específicas de gobernanza adaptadas a sectores concretos.

¿Cómo se articula la gestión de la confianza en el espacio de datos?



3.11. Soluciones y/o servicios que se utilizan para el control de la ejecución de las políticas de uso de datos

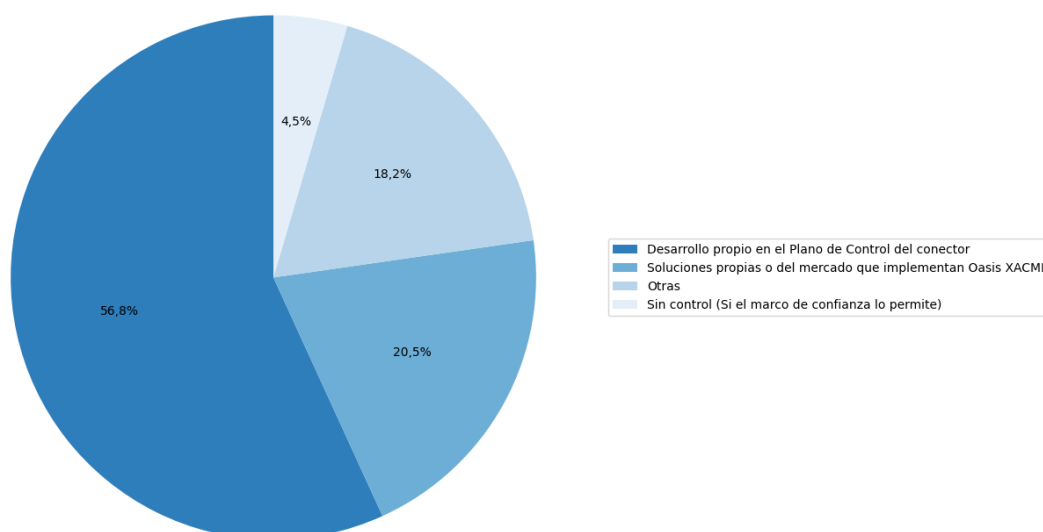
La mayoría de los proyectos, 56,82%, implementa un **desarrollo propio en el Plano de Control del conector** para asegurar la ejecución de las políticas de uso de datos. Esta alta proporción evidencia una preferencia por soluciones personalizadas, posiblemente motivada por la necesidad de adaptar los controles a **contextos regulatorios específicos**, sectores técnicos particulares o integraciones con arquitecturas existentes.

Por otro lado, un 20,45% utiliza tecnologías basadas en **OASIS XACML**, un estándar consolidado para la definición, evaluación y aplicación de políticas de acceso. Esta elección refleja una voluntad de **alinearse con modelos interoperables y reutilizables**, lo que puede facilitar la escalabilidad y la integración en entornos federados.

La categoría **Otras**, 18,18%, agrupa enfoques diversos, incluyendo desde productos comerciales hasta soluciones emergentes o marcos propios aún en desarrollo. Esta dispersión sugiere que el ecosistema sigue explorando alternativas técnicas para implementar mecanismos eficaces de gobernanza del uso de datos.

Finalmente, un pequeño porcentaje, 4,55%, declara no contar con un control explícito de las políticas, siempre que **el marco de confianza lo permita**. Este enfoque puede responder a una **fase temprana del proyecto**, a relaciones de confianza preexistentes o a la dependencia de otras capas de seguridad externa.

¿Qué soluciones y/o servicios se utilizan para el control de la ejecución de las políticas de uso de datos?

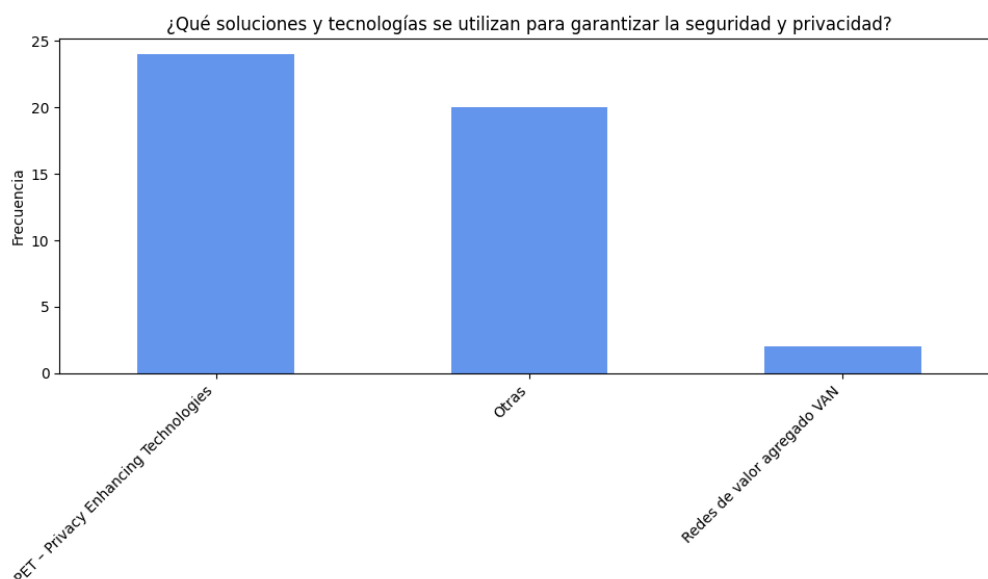


3.12. Soluciones y tecnologías que se utilizan para garantizar la seguridad y privacidad

Las **Privacy Enhancing Technologies (PET)** son la solución más empleada para proteger la seguridad y privacidad de los datos en los espacios analizados, con una adopción del 52,17%. Esta tendencia refleja una clara prioridad por herramientas que permiten **garantizar la confidencialidad, integridad y control sobre los datos personales y sensibles**, incluso durante su procesamiento o compartición.

La categoría **Otras**, 43,48%, sugiere una notable variedad de soluciones, que abarcan desde herramientas comerciales hasta desarrollos propios. Esta diversidad apunta a que muchos proyectos están adoptando **estrategias híbridas o específicas de sector** para abordar retos de seguridad que no siempre pueden resolverse con tecnologías genéricas.

Por último, un 4,35% de los proyectos menciona el uso de **redes de valor agregado (VAN)**, lo que indica su aplicación en escenarios donde se requiere una **infraestructura de comunicación robusta y de confianza** entre múltiples partes, especialmente en sectores como logística, administración electrónica o industria.



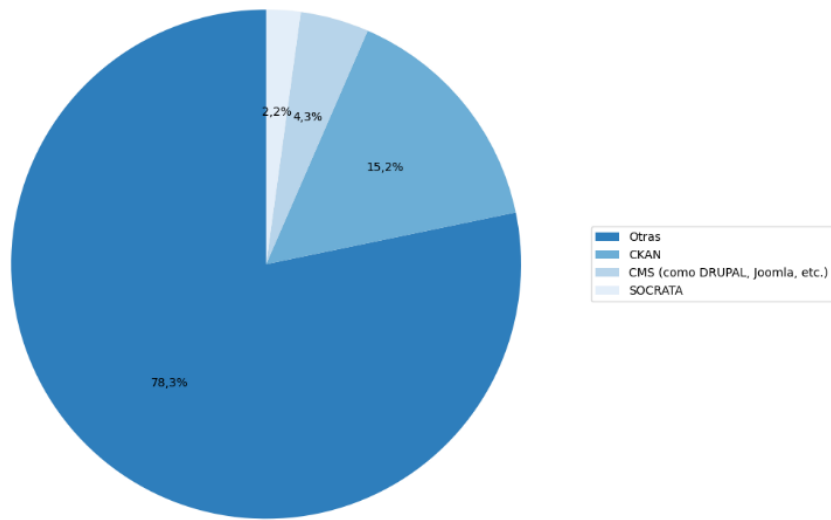
3.13. Herramientas y/o servicios que se utilizan para la catalogación de productos de datos, servicios y otros recursos

Una gran mayoría de los proyectos, 78,26%, reporta el uso de herramientas o servicios **distintos a los más conocidos** para la catalogación de datos, servicios y recursos. Este resultado pone de manifiesto una **alta diversidad tecnológica**, que puede deberse tanto a la adaptación de soluciones existentes a contextos específicos como al empleo de herramientas **propias, sectoriales o aún en fase de evolución**.

La plataforma **CKAN**, reconocida por su orientación a la interoperabilidad y su amplia adopción en portales de datos abiertos, aparece en un 15,22% de los casos. Su presencia refleja una confianza en herramientas consolidadas y alineadas con estándares de publicación y reutilización de datos.

El uso de **CMS** tradicionales como Drupal o Joomla, 4,35%, y de **SOCRATA**, 2,17%, es minoritario, lo que sugiere que se emplean más en proyectos orientados a la **visualización, acceso ciudadano o integración con portales web**, en lugar de como soluciones nativas de catalogación interoperable.

¿Qué herramientas/servicios se utilizan para la catalogación de productos de datos servicios y otros recursos?



5. Anexo A. Cuestionario

Se indican a continuación el detalle de las cuestiones a plantear a los proyectos beneficiarios:

1. Nombre del proyecto/Beneficiario
2. Fase actual del proyecto:
 - a. Diseño/Conceptualización
 - b. Desarrollo/Implementación
 - c. Piloto/Pruebas
 - d. Producción/Operativo
 - e. Mantenimiento/Evolución
3. Nivel de criticidad de los datos:
 - a. Bajo (datos públicos)
 - b. Medio (datos comerciales)
 - c. Alto (datos sensibles/regulados)
 - d. Crítico (datos estratégicos/clasificados)
4. ¿Qué tipo de infraestructura se utiliza en el espacio de datos?
 - a. Cloud
 - b. On-premise
 - c. Híbrido
 - d. Otras (explicar)
5. Si la infraestructura es Cloud, indicar cuál: (Englobar con anterior, para saber cuál es)
 - a. AWS
 - b. Azure
 - c. Google Cloud
 - d. Europea (indicar cuáles: Aire Networks, IONOS, ARUBA, CLEVER, OLH, Interoperability Testbed, Regulatory Sandboxes, Digital Infrastructures de CONNECT, etc.)
 - e. Otros (indicar cuáles)
6. Indicar quiénes son los principales proveedores tecnológicos que os van a ayudar a desarrollar/integrar la solución. Si es un equipo propio indicar “medios propios”. Ordenar por volumen de presupuesto subcontratado. Obligatoria
 - a. Integrador 1
 - b. Integrador 2
 - c. Integrador 3
 - d. Integrador 4
 - e. Medios propios
7. Indicar quiénes son los principales proveedores tecnológicos de software o datos. Ordenar por volumen de presupuesto empleado. Obligatoria

- a. Proveedor 1
 - b. Proveedor 2
 - c. Proveedor 3
 - d. Proveedor 4
8. Arquitecturas de referencia o Building blocks utilizados (Obligatoria)
- a. IDSA
 - b. Gaia-X
 - c. Fiware Reference Model
 - d. DSSC Blueprint Building blocks
 - e. Específica de dominio
 - f. Otros
9. Los datos que se suministran qué modelos y tecnologías utilizan:
- a. Datasets (Indicar desde dónde se suministran: desde el conector, API o frontal web)
 - b. Streaming (Indicar cuáles: KAFKA, LDES, otros)
 - c. Asíncronas (Indicar cuáles: KAFKA, MQTT, RabbitMQ, etc.)
 - d. Otras (indicar cuáles)
10. ¿Qué técnicas de comunicación se utilizan para la transferencia de datos? Obligatoria
- a. Pull
 - b. Push
 - c. Public & Suscribe
 - d. Otros (indicar cuáles)
11. ¿Qué soluciones y/o servicios se utiliza para la transferencia controlada de datos? Obligatoria
- a. EDC, IDSA Connector
 - b. Extensiones y personalizaciones de EDC
 - c. Fiware DSC Connector
 - d. API multipunto
 - e. API punto a punto
 - f. Otros (indicar cuáles)
12. ¿Cómo se articula la gestión de la confianza en el espacio de datos?
- a. Credenciales verificables
 - b. Identificadores descentralizados
 - c. Infraestructura de clave pública.
 - d. Otros (indicar cuáles)
13. ¿Qué soluciones y/o servicios se utilizan para el control de la ejecución de las políticas de uso de datos?
- a. Soluciones propias o del mercado que implementan Oasis XACML (indicar cuáles y qué versiones de XACML. Por ejemplo: MyData Technologies XACML versión 3.0)
 - b. Desarrollo propio en el Plano de Control del conector
 - c. Sin control (Si el marco de confianza lo permite, explicar cómo)
 - d. Otros (indicar cuáles)

14. ¿Qué soluciones y tecnologías se utilizan para garantizar la seguridad y privacidad? Obligatoria
- a. PET – Privacy Enhancing Technologies (concretar cuáles, por ejemplo: encriptación, pseudonimización, ZPK y registros oficiales autoritativos)
 - b. Redes de valor agregado VAN (por ejemplo, VAN EDI, sTesta2, protocolos seguros de internet)
 - c. Otros (indicar cuáles)
15. ¿Qué herramientas/servicios se utilizan para la catalogación de productos de datos, servicios y otros recursos? Obligatoria
- a. CKAN
 - b. SOCRATA
 - c. CMS (como DRUPAL, Joomla, etc.)
 - d. Otros (indicar cuáles)