

PLIEGO DE CONDICIONES QUE REGIRÁN LA CONTRATACIÓN DE UN SERVICIO PARA LA CONSTRUCCIÓN DEL ESPACIO DE DATOS DE LA MÚSICA ESPAÑOLA (KORDAT) Y LA IMPLEMENTACIÓN DE SUS CASOS DE USO, FINANCIADO POR LA UNIÓN EUROPEA – NEXT GENERATION EU, EN EL MARCO DEL PLAN DE RECUPERACIÓN, TRANSFORMACIÓN Y RESILIENCIA

1. INTRODUCCIÓN	3
1.1. Consideraciones previas	3
1.2. Alcance del proyecto KORDAT	4
1.3. Régimen jurídico del contrato	4
2. OBJETO DEL CONTRATO	5
2.1. Actividades a realizar por el adjudicatario	6
Diseño del modelo de datos, modelo de gobernanza, arquitecturas de fuentes, arquitectura técnica, interfaces de usuario y modelo de sostenibilidad del Espacio de Datos	6
Diseño e implementación de la arquitectura del Espacio de Datos	7
Caso de uso 1: Gestión del Dato para la Internacionalización de la Música Española (GDM)	10
Caso de uso 2: Intercambio internacional de derechos conexos	12
Caso de uso 3: Herramientas para la gestión de datos del uso del repertorio del artista (Music Tonic)	12
Gestión de los trabajos	14
2.2 Plazo previsto para la ejecución, hitos y entregables	14
2.3 Informes, entregas y certificaciones	17
3. CAPACIDAD Y SOLVENCIA	17

3.1. Capacidad	17
3.2. Solvencia económica y financiera	17
3.3. Solvencia técnica y profesional	18
3.4. Disposición común a la solvencia económico-financiera y técnica- profesional	19
4. PRESUPUESTO Y VALOR ESTIMADO DEL CONTRATO.....	19
5. PLAZO DE EJECUCIÓN	21
6. PROCEDIMIENTO DE ADJUDICACIÓN, TRAMITACIÓN, LUGAR Y PLAZO DE PRESENTACIÓN DE PROPUESTAS	21
6.1. Publicidad	21
6.2 Forma de presentación de las propuestas.....	21
6.3. Criterios de adjudicación	22
6.4. Adjudicación	23
6.5. Formalización	23
7. MODIFICACIONES	24
8. OBLIGACIONES FORMALES DEL ADJUDICATARIO.....	24
9. CONDICIONES DE FACTURACIÓN Y PAGO.....	25
10. SEGURIDAD, CONFIDENCIALIDAD Y PROTECCIÓN DE DATOS.....	26
11. PROPIEDAD DE LOS TRABAJOS REALIZADOS	27
12. CONTACTO	27
ANEXO I. DECLARACIÓN DE AUSENCIA DE CONFLICTO DE INTERÉS (DACI).....	28
ANEXO II. DECLARACIÓN DE COMPROMISO EN RELACIÓN CON LA EJECUCIÓN DE ACTUACIONES DEL PRTR	30
ANEXO III. ANÁLISIS DE CONTEXTO. ARQUITECTURA DEL ESPACIO DE DATOS	31

1. INTRODUCCIÓN

1.1. Consideraciones previas

El servicio objeto de este contrato forma parte del proyecto “*Construcción del Espacio de Datos de la música española*”, beneficiario de la convocatoria de ayudas en el ámbito de la digitalización, para la transformación digital de los sectores productivos estratégicos mediante la creación de demostradores y casos de uso de Espacios de compartición de Datos, en el marco del Plan de Recuperación, Transformación y Resiliencia -Next Generation EU (Programa de datos Sectoriales, Componente 12, Inversión 1).

El sector musical español y europeo es un pilar fundamental tanto para la cultura como para la economía. La industria musical española ha experimentado una profunda transformación digital en la última década, lo que subraya la necesidad de una mayor interoperabilidad, transparencia y seguridad en la gestión de sus datos. A pesar de su vitalidad, el sector se enfrenta a desafíos significativos: los datos están dispersos y son difícilmente accesibles u operables fuera del contexto de las organizaciones que los custodian, lo que limita la transparencia, el análisis profundo y la toma de decisiones informada.

Es en este contexto que la creación del Espacio de Datos de la Música Española, con el nombre de KORDAT, emerge como una oportunidad transformadora y una necesidad estratégica. Este proyecto pionero en España tiene como objetivo establecer una infraestructura de datos compartida que facilite la interoperabilidad, el análisis avanzado y la generación de valor para todos los actores de la cadena de valor musical. KORDAT se concibe como un ecosistema donde la compartición voluntaria de datos se realiza en un entorno de soberanía, confianza y seguridad, basado en mecanismos integrados de gobernanza, organizativos, normativos y técnicos. Esta iniciativa está promovida por la Sociedad de Artistas Intérpretes o Ejecutantes (AIE) y se enmarca en el impulso de los Espacios Europeos de Datos Sectoriales, cofinanciado por los fondos Next Generation EU del Ministerio para la Transformación Digital.

La conveniencia de este espacio de datos se traduce en beneficios tangibles para el ecosistema musical en general, ya que el objetivo de KORDAT es convertirse en un catalizador de la innovación empresarial, fomentando la digitalización avanzada basada en datos, y abriendo nuevas oportunidades de negocio más allá de las prácticas actuales de la industria musical.

Para ello, la arquitectura de KORDAT incorporará las últimas tecnologías, y adoptará un modelo de gobernanza robusto y transparente, con canales de comunicación y servicios específicos para garantizar la accesibilidad y el máximo valor para todos, y

se promoverá una semántica y un vocabulario comunes para estandarizar el intercambio de información.

El diseño de la plataforma será interoperable con los Espacios de Datos Comunes Europeos, en línea con iniciativas como Gaia-X e IDS-RAM, y fomentará el uso de fuentes de datos abiertos.

En resumen, el Espacio de Datos de la Música Española representa una apuesta decidida por la innovación y la colaboración, impulsando la transformación de un sector fragmentado desde el punto de vista de los datos hacia un ecosistema conectado y transparente.

1.2. Alcance del proyecto KORDAT

El objetivo principal de KORDAT es el desarrollo de un Espacio de Datos común para la Música Española, contando con AIE como promotor inicial del proyecto. La plataforma estará sustentada en la utilización de tecnologías y soluciones cloud para intercambiar, capturar, consolidar, procesar y analizar múltiples fuentes de datos. El resultado final será, por un lado, la arquitectura e implementación del Espacio de Datos, y por otro, herramientas de apoyo concretas para la implementación de cada caso de uso.

Los casos de uso inicialmente contemplados son los siguientes:

- Gestión del dato para la Internacionalización de la Música Española (GDM).
- Intercambio internacional de derechos conexos (Derechos conexos).
- Herramienta para gestión de datos del uso del repertorio del artista (Music Tonic).

Los resultados esperados para los casos relacionados con la internacionalización y el repertorio será la creación de los repositorios y catálogos de información que faciliten la integración de los datos en informes detallados y cuadros de mando interactivos, configurados para convertirse en una herramienta de apoyo a la toma de decisiones estratégicas en el sector musical. Para el caso de los derechos conexos, el objetivo es que KORDAT contribuya a optimizar el intercambio de derechos entre sociedades de gestión, entre otros beneficios.

1.3. Régimen jurídico del contrato

La entidad contratante es la asociación "*Artistas Intérpretes o Ejecutantes, EGDPI*" (AIE), que es una entidad de gestión de derechos de propiedad intelectual de los artistas intérpretes o ejecutantes y de sus derechohabientes, autorizada de conformidad con el art. 147 de la Ley de Propiedad Intelectual (Real Decreto Legislativo 1/1996, de 12 de abril, por el que se aprueba el texto refundido de la Ley de Propiedad Intelectual).

El presente contrato tiene naturaleza privada, y se regirá por las normas del ordenamiento jurídico privado, en lo no previsto en estos pliegos, sin perjuicio de las normas aplicables en función de la cofinanciación con cargo a fondos europeos.

Para cuantas cuestiones pudieran plantearse entre los contratistas y la Entidad, y entre aquéllos como tales, con renuncia expresa a cualquier otro fuero que pudiera corresponderles, unos y otra declaran aplicable la Ley española y se someten expresamente a los Juzgados y Tribunales del domicilio social de la Entidad. Sin perjuicio de lo establecido en el párrafo anterior, cualquier controversia que pudiera surgir entre la Entidad y los contratistas, se someterá antes de su planteamiento judicial a un trámite previo de conciliación.

2. OBJETO DEL CONTRATO

El presente contrato, alineado con la memoria presentada por AIE en la Orden TDF/1461/2023, de 29 de diciembre, por la que se aprobaban las bases reguladoras en el ámbito de la digitalización, para la transformación digital de los sectores productivos estratégicos mediante la creación de demostradores y casos de uso de Espacios de Compartición de Datos, en el marco del Plan de Recuperación, Transformación y Resiliencia-Next Generation EU (Programa Espacios de Datos Sectoriales), tiene como objetivos principales:

- **Diseñar el modelo de datos, modelo de gobernanza, arquitecturas de fuentes, arquitectura técnica, interfaces de usuario y modelo de sostenibilidad de KORDAT**, así como **definir de manera detallada los casos de uso** especificados en el proyecto. Estos trabajos, considerados dentro del **Paquete de Trabajo 1 “Especificaciones, diseño de la arquitectura y definición detallada de los casos de uso”**, deberán partir necesariamente de los resultados obtenidos del análisis de contexto preliminar ya completado.
- **Desarrollar la infraestructura IT sobre la que se asienta KORDAT** y que permitirá el intercambio transparente, soberano y seguro de datos entre diferentes actores independientes de la cadena de valor de la música, lo que incluye tanto los desarrollos informáticos necesarios como la configuración e implementación de la infraestructura cloud que soporte los desarrollos. Estos trabajos, definidos dentro del **Paquete de Trabajo 2 “Implementación de la plataforma ED”**, deberán tener en cuenta las arquitecturas de referencia de espacios de datos a nivel internacional y, en particular, las que se deriven de los trabajos del UNE CT71/SC43 para la definición de un marco de referencia técnico y de interoperabilidad, sirviendo de contribución nacional al CEN/CENELEC e ISO.
- **Validar la arquitectura de referencia adoptada mediante la implementación de tres casos de uso** relacionados con la internacionalización de la música, el intercambio internacional de derechos conexos, así como la

gestión de datos de los artistas. Estos trabajos son parte de las tareas detalladas en el **Paquete de Trabajo 3 “Validación del ED a través del desarrollo de los casos de uso”**.

Para todo ello, AIE requiere la asistencia de un proveedor con capacidad para prestar los servicios de consultoría estratégica, legal y tecnológica, y afrontar los desarrollos informáticos que requieren los trabajos objeto de proyecto KORDAT.

2.1. Actividades a realizar por el adjudicatario

En este apartado se especifican los requisitos mínimos que deberán cumplir los trabajos que forman parte del objeto de la licitación, si bien los mismos podrán ser objeto de mejoras adicionales propuestas.

Las propuestas que ofrezcan características inferiores a las requeridas en el presente pliego podrán ser clasificadas como no conformes y por tanto incurrir en causa de exclusión del proceso de adjudicación.

En este sentido, el licitador podrá proponer, dentro de su oferta, prestaciones que mejoren o amplíen las requeridas en el presente pliego, siempre que dichas mejoras no supongan un incremento del presupuesto total establecido como máximo en esta licitación. Estas prestaciones adicionales deberán estar debidamente justificadas, detallando su valor añadido respecto a los requisitos mínimos exigidos, y serán tenidas en cuenta en el proceso de valoración conforme a los criterios establecidos

Diseño del modelo de datos, modelo de gobernanza, arquitecturas de fuentes, arquitectura técnica, interfaces de usuario y modelo de sostenibilidad del Espacio de Datos

Los paquetes de trabajo, actuaciones y entregables detallados a lo largo del presente pliego se corresponden con la memoria técnica presentada por la entidad y aprobada por el Ministerio de Cultura en el marco de la Orden TDF/1461/2023, de 29 de diciembre, por la que se conceden ayudas destinadas a la creación de demostradores y casos de uso de Espacios de Compartición de Datos, en el contexto del Programa de Espacios de Datos Sectoriales del Plan de Recuperación, Transformación y Resiliencia – Next Generation EU.

Tomando en consideración los resultados obtenidos en la actividad¹ de "**Análisis de contexto**", incluida en el paquete de trabajo PT1, y cuyo resultado se anexa al

¹ Existe una pequeña divergencia en la codificación de esta actividad, ya que en la memoria presentada figura como "A1.2", mientras que en la notificación de resolución recibida se denomina "A.9" del PT1.

presente pliego como Anexo III, el objetivo general es completar el diseño de los elementos constituyentes del Espacio de Datos y que se detallan en las actividades correspondientes del PT1:

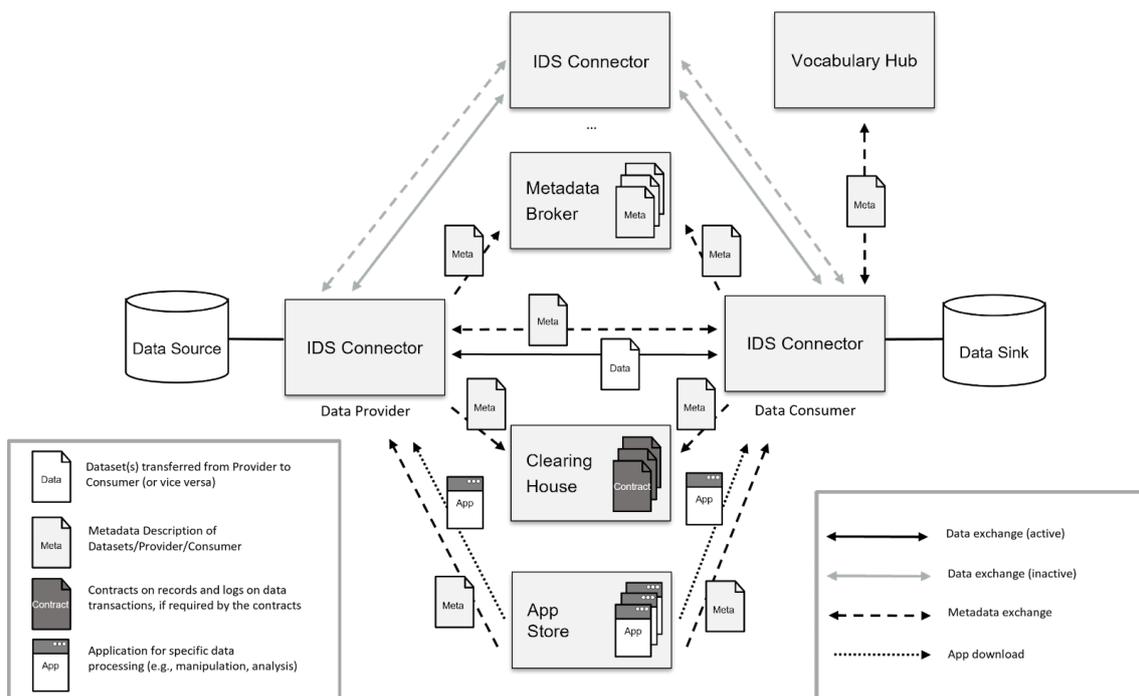
- **A1.2. Diseño del modelo de datos:** definición de la taxonomía del ED, adecuación a estándares, normalización.
- **A1.3. Diseño de la arquitectura de fuentes:** catalogación de fuentes, modelos de datos, trazabilidad y mapeado de datos
- **A1.4. Diseño de la arquitectura técnica:** identificación de módulos técnicos, selección de tecnologías, integración, requisitos técnicos.
- **A1.5. Diseño de interfaces de usuario:** identificación de canales de explotación, diseño de interacción, diseño visual
- **A1.6. Definición detallada de casos de uso:** diseño de detalle de los casos de USO
- **A1.7. Diseño del modelo de sostenibilidad:** modelo de gobernanza, modelo de negocio, oferta de servicios y usos previstos.

Para la correcta ejecución de los trabajos previstos en estas actividades, los licitadores deberán demostrar que su equipo cuenta con las capacidades adecuadas en materia de diseño de modelos de negocio, modelos de gobernanza y aspectos legales. En este sentido, se valorarán los perfiles no técnicos aportados en el equipo de trabajo, o bien los acuerdos con empresas colaboradoras especializadas que puedan complementar la labor técnica.

Diseño e implementación de la arquitectura del Espacio de Datos

Los resultados obtenidos en la actividad de "**Análisis de contexto**", que se adjuntan como Anexo III en el presente pliego, sugieren considerar la arquitectura IDS-RAM 4.0, propuesta por IDSA (International Data Spaces Association), como el marco de trabajo de referencia más adecuado para el desarrollo del Espacio de Datos de la Música.

Este modelo define un conjunto de componentes fundamentales que permiten el intercambio seguro, confiable y soberano de datos entre organizaciones. Estos componentes trabajan en conjunto para garantizar la autenticación de los participantes, la protección de los datos, la aplicación de políticas de uso y la interoperabilidad técnica y semántica dentro del ecosistema IDS (ver figura).



Dentro de la arquitectura IDS RAM podemos distinguir diversos componentes, algunos obligatorios, para la implementación de un espacio de datos mínimo viable, y otros opcionales que llevan a funcionalidades adicionales. Las piezas imprescindibles incluyen:

- **IDS Connector:** obligatorio para todos los participantes. Es el núcleo del ecosistema IDS. Sin él, no se puede intercambiar datos ni aplicar políticas de uso. Actúa como una puerta de enlace que permite a las organizaciones conectar sus sistemas existentes al ecosistema IDS, facilitando el intercambio de datos de manera segura y controlada. El conector gestiona aspectos como la autenticación, autorización, gestión de metadatos y aplicación de políticas de uso de datos.
- **Identity Provider (IdP):** obligatorio a nivel de sistema. Todos los conectores deben autenticar su identidad mediante un proveedor de identidad certificado por IDSA.
- **Policy Enforcement Mechanism (en el conector):** obligatorio dentro del conector. Debe existir un mecanismo para aplicar y respetar las políticas de uso de datos acordadas.

Los siguientes componentes no son estrictamente necesarios, pero recomendables para enriquecer la funcionalidad, la trazabilidad o la interoperabilidad:

- **Metadata Broker:** opcional pero altamente recomendado. Facilita el descubrimiento de conectores y ofertas de datos. No es obligatorio si los conectores se conocen directamente.

- **Clearing House:** opcional. Aporta trazabilidad y auditoría, útil en entornos con requisitos regulatorios o de confianza reforzada.
- **App Store:** opcional. Permite ejecutar apps certificadas sobre los datos. Útil si se requiere procesamiento dentro del conector.
- **Vocabulary Provider / Ontology Hub:** opcional. Soporta interoperabilidad semántica, muy útil en contextos intersectoriales o multilingües.
- **Certification Service:** opcional (pero necesario para participantes certificados). Aunque no todos los participantes necesitan certificarse, es obligatorio para quienes operan bajo requisitos formales de seguridad y calidad.

Por otro lado, junto con las arquitecturas de referencia de espacios de datos a nivel internacional, el UNE CT71/SC43 trabaja en la definición de un marco de referencia técnico y de interoperabilidad, sirviendo de contribución nacional al CEN/CENELEC e ISO, que será previsiblemente publicado en julio de 2025.

Con todo ello, los licitadores deberán presentar una propuesta que contemple la realización de las siguientes actividades del proyecto, y que forman parte del PT2:

- **A2.1. Módulo de interoperabilidad de datos:** modelos de datos, intercambio de datos, origen y trazabilidad.
- **A2.2. Módulo de Soberanía de datos y confianza:** políticas de acceso y uso y control, gestión de identidad, confianza.
- **A2.3. Módulo de Creación de Valor del Dato:** descripción de ofertas, servicios y datos, publicación y descubrimiento, marketplace.
- **A2.4. Módulo de Negocio:** desarrollo de modelo de negocio, desarrollo de caso de uso, desarrollo de producto del dato, intermediarios del ED.
- **A2.5. Módulo de Gobernanza:** gobernanza organizaciones, gobernanza de compartición de datos.
- **A2.6. Módulo legal:** cumplimiento regulatorio, marco contractual.
- **A2.7. Módulo de infraestructura en cloud:** arquitectura en nube que soporte el desarrollo del proyecto.

Para la correcta ejecución de los trabajos previstos en las **actividades A2.5 y A2.6**, los licitadores deberán demostrar que su equipo cuenta con las capacidades adecuadas en materia de diseño de modelos de negocio, modelos de gobernanza y aspectos legales. En este sentido, se valorarán los perfiles no técnicos aportados en el equipo de trabajo, o bien los acuerdos con empresas colaboradoras especializadas que puedan complementar la labor técnica.

En cuanto a la **actividad A2.7**, los licitadores deberán proveer una solución tipo “llave en mano” que contemple tanto la implementación de la arquitectura del espacio de datos en el proveedor cloud elegido, preferiblemente AWS, como los consumos de datos y otros servicios que pudieran derivarse del funcionamiento del espacio de datos.

Es fundamental destacar en relación con la **actividad A2.7**, que la **plataforma residirá en un entorno tecnológico específico e independiente de los sistemas de AIE**. En este sentido, la plataforma debe gestionar la información proveniente de AIE del mismo modo que lo haría para cualquier otro agente participante que quisiera integrar parte de sus datos en el espacio. La plataforma, por tanto, estará dimensionada para la entrada de diversas fuentes de datos y participantes. Como primer escenario, para disminuir las posibles barreras de entrada en el espacio, AIE proveerá al resto de participantes del espacio con las herramientas necesarias para operar, tales como el conector y un espacio de almacenamiento, lo que deberá ser tenido en cuenta por parte de los licitadores.

Por tanto, se establecen como requisitos mínimos de la Actividad 2 de este pliego:

- El alineamiento con la arquitectura de referencia IDS-RAM 4.0 y sus principios fundamentales, y con los trabajos derivados del UNE CT71/SC43 para la definición de un marco de referencia técnico y de interoperabilidad, que serán previsiblemente publicados en julio de 2025
- El desarrollo de un espacio de datos mínimamente viable, según IDS. Cualquier componente o funcionalidad adicional será valorada positivamente en base al valor aportado.
- La aportación a los participantes de las herramientas básicas para operar en el espacio.

Caso de uso 1: Gestión del Dato para la Internacionalización de la Música Española (GDM)

El caso de uso GDM (Gestión de Datos Musicales) tiene como objetivo la construcción de una herramienta que ofrezca soporte al análisis de situación y diseño de iniciativas para la internacionalización y promoción de la música española. Para ello, la herramienta se concibe como un portal que permita visualizar los datos de diferentes fuentes en un mismo entorno, en el que los agentes participantes aportan su información, junto con otros datos extraídos de fuentes complementarias.

Los datos que se gestionarán en este caso de uso son potencialmente muy diversos. En primer lugar, se incluyen datos internos proporcionados por AIE, como el monitoreo musical, el repertorio y el lineup. Además, los usuarios registrados podrán aportar información mediante la carga de archivos en formato CSV. Este flujo de datos se complementará con información adicional obtenida a través de APIs e integraciones con fuentes externas, como plataformas de streaming, redes sociales y bases de datos públicas, entre ellas las del Instituto Nacional de Estadística (INE).

Para ello, se establecen como requisitos mínimos de esta actividad, **denominada A3.2:**

- Implementación del caso de uso sobre la infraestructura del espacio de datos.
- Ingesta de datos:
 - La extracción y carga de datos desde las diferentes fuentes y participantes al espacio de datos.
 - El procesamiento de datos que faciliten su limpieza y anonimizado, permitiendo su transformación y disponibilización.
 - El enriquecimiento de la información a través de agregaciones y el cálculo de estadísticas, métricas e indicadores que describan y resuman los datos de manera significativa.
 - La facilitación del acceso a datos estructurados a través de APIs, con preferencia por formatos CSV o JSON, y con la documentación adecuada de los conectores.
- Almacenamiento de datos.
- Configuración del modelo de datos de GDM:
 - Estar alineado con la terminología estándar utilizada tanto en AIE como en las diferentes sociedades de gestión internacionales con las que colabora.
 - Estar diseñado para cumplir con los informes y entregables previstos.
 - Incluir un “modelo de crecimiento” para acomodar datos futuros.
 - Permitir comparaciones entre diferentes años y países, y ser escalable para adaptarse a la información de fases sucesivas (en la medida en que se pueda prever).
 - Implementar un sistema de 'curado' que garantice que solo los datos que cumplan con los estándares establecidos sean incorporados en la plataforma. Esto asegurará que la información irrelevante o que pueda distorsionar los análisis sea identificada y descartada, manteniendo así la calidad y relevancia de los datos para el análisis efectivo y la toma de decisiones informadas.
- Integrar y armonizar los datos obtenidos de las actividades anteriores.
- Analizar los datos para desarrollar indicadores estratégicos que sean directamente aplicables y aporten valor significativo a los objetivos establecidos en este pliego.
- Diseñar y producir informes detallados y cuadros de mando dinámicos que serán accesibles a través de un portal web intuitivo y específicamente creado para este propósito.

Caso de uso 2: Intercambio internacional de derechos conexos

Este caso de uso está orientado a facilitar el intercambio de información entre sociedades de gestión de diferentes países, que típicamente cuenta con un acuerdo de reciprocidad para intercambiar la recaudación de derechos de artistas. Mediante estos acuerdos, las sociedades trasladan a su homóloga en el país correspondiente la recaudación nacional que corresponde a artistas cuya representación corresponde a esa sociedad. De este modo, existe un flujo económico permanente entre sociedades, que periódicamente liquidan derechos entre sí, y que posteriormente cada sociedad nacional hace efectivos a sus representados.

En este contexto, el objetivo es crear un espacio que permita prestar un conjunto de servicios compartidos, que van desde el intercambio y almacenamiento de datos al procesado de datos (adaptación de formatos, por ejemplo) para reforzar la operación local de cada sociedad con un entorno común.

El valor que proporciona este caso es una mejora de la gestión de derechos conexos a nivel internacional, beneficiando en último lugar al colectivo de artistas que AIE representa, pero también al representado por cualquier otra sociedad de gestión participante. Además, este caso contribuye al desarrollo y adopción de buenas prácticas de gobierno corporativo, principalmente en lo referente a la transparencia en el funcionamiento de las sociedades y su obligación de reporte hacia el colectivo representado y los organismos de supervisión y control de cada país.

Para ello, se establecen como requisitos mínimos de esta actividad, **denominada A3.4:**

- Acordar un formato estándar para el intercambio de información.
- Implementar mecanismos de ingesta y validación de datos para garantizar la calidad y consistencia.
- Desarrollar frontend y backend para la interfaz de usuario y la operativa.
- Proporcionar capacidades analíticas para el seguimiento detallado de operaciones y derechos.
- En concreto, se define como punto de partida de este caso el intercambio de datos sobre liquidaciones económicas entre entidades, siguiendo el estándar DDEX.

Caso de uso 3: Herramientas para la gestión de datos del uso del repertorio del artista (Music Tonic)

Con la consolidación de las plataformas digitales como principal entorno de distribución y consumo musical —conviviendo aún con los medios tradicionales analógicos—, se hace evidente la dificultad que enfrentan los artistas y creadores

para gestionar y analizar la enorme cantidad de información relacionada con el uso de su repertorio y la monetización asociada. Esta información les llega habitualmente de forma desestructurada, heterogénea y dispersa, procedente de discográficas, distribuidoras o agregadores; de entidades de gestión de derechos de propiedad intelectual (tanto de autor como de intérprete); y de una amplia variedad de plataformas digitales internacionales (como Spotify, Deezer, Tidal, Amazon, Apple, Netflix, HBO, entre otras), sin que exista una relación ni integración estructurada entre todas estas fuentes.

Music Tonic tiene como objetivo dotar al artista de una plataforma digital accesible online, desde cualquier lugar y dispositivo, con la que el artista/creador pueda tomar control sobre el uso de sus datos, facilitando así su explotación directa o bien a través de plataformas y servicios de terceros que puedan estar autorizados para acceder a dichos datos.

En este sentido, el espacio de datos ofrecerá las funcionalidades técnicas necesarias, estableciéndose como requisitos mínimos de esta actividad, **denominada A3.6**, los siguientes:

- Implementación del caso de uso sobre la infraestructura del espacio de datos.
- Ingesta de datos:
 - La extracción y carga de datos desde las diferentes fuentes y participantes al espacio de datos.
 - El procesamiento de datos que faciliten su limpieza y anonimizado, permitiendo su transformación y disponibilización.
 - El enriquecimiento de la información a través de agregaciones y el cálculo de estadísticas, métricas e indicadores que describan y resuman los datos de manera significativa.
 - La facilitación del acceso a datos estructurados a través de APIs, con preferencia por formatos CSV o JSON, y con la documentación adecuada de los conectores.
- Almacenamiento de datos.
- Configuración del modelo de datos Music Tonic:
 - Estar alineado con la terminología estándar utilizada tanto en AIE como en las diferentes sociedades de gestión internacionales con las que colabora.
 - Estar diseñado para cumplir con los informes y entregables previstos.
 - Incluir un “modelo de crecimiento” para acomodar datos futuros.
 - Implementar un sistema de 'curado' que garantice que solo los datos que cumplan con los estándares establecidos sean incorporados en la plataforma. Esto asegurará que la información irrelevante o que pueda distorsionar los análisis sea identificada y descartada, manteniendo así la calidad y relevancia de los datos para el análisis efectivo y la toma de decisiones informadas.

- Integrar y armonizar los datos obtenidos de las actividades anteriores.
- Analizar los datos para desarrollar indicadores estratégicos que sean directamente aplicables y aporten valor significativo a los objetivos establecidos en este pliego.
- Diseñar y producir informes detallados y cuadros de mando dinámicos que serán accesibles a través de un portal web intuitivo y específicamente creado para este propósito.

Gestión de los trabajos

El adjudicatario deberá incluir en su propuesta un modelo de gestión y buen gobierno de los trabajos contemplados en el presente pliego que contribuya a su correcto desarrollo y la consecución de los objetivos y actividades descritas anteriormente.

2.2 Plazo previsto para la ejecución, hitos y entregables

El plazo máximo de ejecución de los trabajos detallados en el presente pliego será **desde la fecha de adjudicación hasta la fecha de finalización del proyecto KORDAT, prevista para el 15 de junio de 2026** y que se establecerá en el contrato. El incumplimiento de este plazo implicará la resolución del contrato.

Los **entregables** que serán de necesario cumplimiento por parte de los licitadores se presentan en la siguiente tabla:

Código	Entregable	Descripción	Clasificación	Paquete de Trabajo
E1.1	Diseño técnico del ED	Descripción técnica del diseño propuesto para el ED, incluyendo los building blocks y componentes propuestos en la solución. Se incluirán los requisitos a cumplir por el sistema ED, el modo de verificación del requisito y un KPI cuantificable, cuando sea posible.	Documentación	PT1
E1.2	Diseño funcional del ED	Descripción de la operativa propuesta para el funcionamiento del ED.	Documentación	PT1
E1.3	Análisis detallado de los casos de uso	Elaboración de un documento con la identificación de los diferentes casos de uso que se incorporarán en la plataforma digital, así como un análisis detallado de aspectos clave como los participantes del caso de uso, su rol, las fuentes de datos requeridas o las funciones esperadas.	Documentación	PT1
E1.4	Modelo de gobernanza	Documento que describe en detalle el modelo operativo y de gestión del ED. Incluye detalle de roles, procedimientos, herramientas y políticas de gestión.	Documentación	PT1
E1.5	Estrategia comercial	Documento que define la estrategia de comercialización y evolución de negocio del ED. Incluye un plan de crecimiento, vías de generación de ingresos, modelo de venta y objetivos comerciales.	Documentación	PT1
E2.1	Desarrollo de la plataforma v intermedia	Descripción del desarrollo del ED, cubriendo las diferentes piezas que lo componen (building blocks y componentes), así como la metodología de trabajo y el enfoque aplicados. Versión intermedia.	Software	PT2
E2.2	Desarrollo de la plataforma v final	Descripción del desarrollo del ED, cubriendo las diferentes piezas que lo componen (building blocks y componentes), así como la metodología de trabajo y el enfoque aplicados. Versión final.	Documentación	PT2
E2.3	Prototipo de la plataforma del ED	Maqueta del ED desarrollada que describe los flujos más relevantes del mismo para permitir validar con usuarios.	Software	PT2

Código	Entregable	Descripción	Clasificación	Paquete de Trabajo
E3.1	Informe de implementación de casos de uso	Descripción del despliegue de los casos de uso propuestos en el proyecto, incluyendo el itinerario completo del dato desde su ingesta hasta su consumo para el desarrollo de las diferentes funcionalidades propuestas para cada caso.	Documentación	PT3
E3.2	Prototipos de casos de uso	Pantallazos del prototipo desarrollado para cada caso de uso.	Software	PT3

Tabla 1. Entregables de paquetes de trabajo del proyecto

Los licitadores deberán presentar una propuesta detallada de planificación temporal de las diferentes actividades, hitos y entregables, que será finalmente validada y ratificada a la firma del contrato con el equipo adjudicatario.

2.3 Informes, entregas y certificaciones

A la conclusión del proyecto, el adjudicatario presentará los entregables finales que serán validados o certificados por el equipo de Transformación Digital de AIE. Asimismo, el adjudicatario se compromete a colaborar activamente en las tareas de justificación de los trabajos realizados ante cualquier requerimiento del Ministerio de Transformación Digital.

3. CAPACIDAD Y SOLVENCIA

3.1. Capacidad

Podrán contratar las personas naturales o jurídicas, españolas o extranjeras, con plena capacidad de obrar y que acrediten solvencia económica, financiera, técnica y profesional, tal como se indica en los apartados 3.2 y 3.3 del presente documento.

Los licitadores deberán tener una finalidad o actividad que tenga relación directa con el objeto del contrato, según resulte de sus respectivos estatutos o reglas fundacionales, así como disponer de una organización con elementos personales y materiales suficientes para la debida ejecución del contrato.

En su propuesta, los licitadores deberán explicitar que *"este proyecto será financiado por la Unión Europea-Next Generation EU, al amparo de la convocatoria de ayudas Espacios de Datos Sectoriales 2023 para la creación de demostradores y casos de uso de espacios de compartición de datos, y de conformidad con lo dispuesto en el artículo 34 de la Orden TDF/1461/2023, de 29 de diciembre, por la que se aprueban las bases reguladoras y se efectúa la convocatoria para la concesión de ayudas, en el ámbito de la digitalización, para la transformación digital de los sectores productivos estratégicos mediante la creación de demostradores y casos de uso de Espacios de Compartición de Datos, en el marco del Plan de Recuperación, Transformación y Resiliencia-Next Generation EU (Programa Espacios de Datos Sectoriales), modificada por la Orden TDF/294/2024, de 4 de abril"*.

3.2. Solvencia económica y financiera

La solvencia económica y financiera se acreditará mediante:

- Si se trata de una **persona jurídica**, cualquiera de los siguientes medios:

- Documentación acreditativa de la experiencia en la prestación de servicios similares al objeto de contratación, en alguno de los tres últimos ejercicios, junto con el importe de la cifra de negocio asociada a la prestación de dichos servicios, IVA incluido.
- Declaraciones de entidades financieras, a través de un informe de la entidad financiera en el que la entidad financiera se pronuncie expresamente sobre la solvencia económica y financiera del licitador para llevar a cabo la ejecución del objeto del contrato.
- En el caso de una sociedad mercantil, se presentarán las cuentas anuales de los 3 últimos ejercicios presentadas en el Registro Mercantil, y los modelos 200 de declaración-liquidación del Impuesto de Sociedades de los 3 últimos ejercicios.
- En el caso de **personas físicas**:
 - Si tributa en el IRPF mediante estimación directa:
 - Libro registro de ventas e ingresos de los 3 últimos años
 - Modelo 390 del IVA de los 3 últimos años
 - Modelo 100 de ingreso o devolución de la declaración del IRPF de los 3 últimos años.
 - Si tributa en el IRPF mediante estimación objetiva:
 - Libro registro de ventas e ingresos de los 3 últimos años y/o modelo 100 de los 3 últimos años.
 - Declaración responsable firmada por el representante legal que acredite que el libro de registro de ventas e ingresos refleja la imagen fiel de la actividad económica de la empresa. En caso de no disponer de este libro, la declaración debe acreditar que el rendimiento neto de la actividad refleja la imagen fiel de la actividad económica de la empresa.

3.3. Solvencia técnica y profesional

Se acreditará mediante la relación de los principales servicios o trabajos similares realizados en los últimos tres años, incluyendo importe, fechas y destinatario. Se valorará la experiencia acumulada en proyectos de naturaleza similar, particularmente en el ámbito de la industria musical, la gestión de proyectos financiados con fondos públicos o la participación en iniciativas de apoyo a la transformación digital de organizaciones.

3.4. Disposición común a la solvencia económico-financiera y técnica- profesional

El licitador podrá presentar simplemente una declaración responsable en el momento de presentación de su oferta, haciendo constar que dispone de la solvencia exigida, sin perjuicio de que, en todo caso, el adjudicatario deberá presentar la documentación exigida en los puntos 3.2 y 3.3.

4. PRESUPUESTO Y VALOR ESTIMADO DEL CONTRATO

El importe máximo del contrato para los trabajos descritos anteriormente será de seiscientos siete mil novecientos ochenta y cinco euros (607.985 Euros) (IVA no incluido), cuyo desglose por actividad se muestra a continuación:

Código	Actividad	Paquete de Trabajo	Presupuesto máximo Financiable (€)
A1.2	Diseño del modelo de datos	PT1 - Especificaciones, diseño de la arquitectura y definición de los casos de uso	8.000,00
A1.3	Diseño de la arquitectura de fuentes	PT1 - Especificaciones, diseño de la arquitectura y definición de los casos de uso	5.000,00
A1.4	Diseño de la arquitectura técnica	PT1 - Especificaciones, diseño de la arquitectura y definición de los casos de uso	10.000,00
A1.5	Diseño de interfaces de usuario	PT1 - Especificaciones, diseño de la arquitectura y definición de los casos de uso	8.000,00
A1.6	Definición detallada de casos de uso	PT1 - Especificaciones, diseño de la arquitectura y definición de los casos de uso	5.000,00
A1.7	Diseño del modelo de sostenibilidad	PT1 - Especificaciones, diseño de la arquitectura y definición de los casos de uso	10.000,00
A2.1	Módulo de interoperabilidad de datos	PT2 - Implementación de la plataforma ED	80.000,00

Código	Actividad	Paquete de Trabajo	Presupuesto máximo Financiable (€)
A2.2	Módulo de Soberanía de datos y confianza	PT2 - Implementación de la plataforma ED	70.000,00
A2.3	Módulo de Creación de Valor del Dato	PT2 - Implementación de la plataforma ED	65.000,00
A2.4	Módulo de Negocio	PT2 - Implementación de la plataforma ED	25.000,00
A2.5	Módulo de Gobernanza	PT2 - Implementación de la plataforma ED	44.000,00
A2.6	Módulo legal	PT2 - Implementación de la plataforma ED	40.000,00
A2.7	Módulo de infraestructura en cloud	PT2 - Implementación de la plataforma ED	97.985,00
A3.2	Desarrollo del Caso de Uso 1: GDM	PT3 - Validación del ED a través del desarrollo de los casos de uso	40.000,00
A3.4	Desarrollo del Caso de Uso 2: Derechos de Reproducción	PT3 - Validación del ED a través del desarrollo de los casos de uso	50.000,00
A3.6	Desarrollo del Caso de Uso 3: Music Tonic	PT3 - Validación del ED a través del desarrollo de los casos de uso	50.000,00
		TOTAL	607.985,00 €

Tabla 2. Detalle de presupuesto máximo financiable por actividad

En el precio del contrato se considerarán incluidos los tributos, tasas y cánones de cualquier índole que sean de aplicación excepto el IVA.

5. PLAZO DE EJECUCIÓN

El plazo máximo de ejecución de los trabajos detallados en el presente pliego será **desde la fecha de adjudicación hasta la fecha de finalización del proyecto KORDAT, prevista para el 15 de junio de 2026**, que se establecerá en el contrato. El incumplimiento de este plazo implicará la resolución del contrato.

El incumplimiento del plazo de ejecución implicará la resolución del contrato, obligándose el adjudicatario a devolver, en su caso, las cantidades recibidas en concepto de pagos parciales, sin perjuicio de la indemnización que corresponda por daños y perjuicios.

6. PROCEDIMIENTO DE ADJUDICACIÓN, TRAMITACIÓN, LUGAR Y PLAZO DE PRESENTACIÓN DE PROPUESTAS

6.1. Publicidad

La presente licitación se tramita mediante procedimiento abierto, garantizando los principios de publicidad, concurrencia y transparencia. El anuncio se publicará en la página web de AIE, a fin de facilitar el acceso a toda la información relativa al proceso. Asimismo, se cursará invitación directa a un mínimo de tres empresas con experiencia contrastada en la materia objeto del contrato, con el objetivo de asegurar una participación representativa y competitiva en el procedimiento.

Desde la publicación del anuncio, los licitadores que deseen participar en este procedimiento **deberán presentar sus propuestas de forma telemática** en la dirección de correo secregen@aie.es en un **plazo máximo de 15 días naturales a partir del día siguiente al de la publicación del anuncio**.

6.2 Forma de presentación de las propuestas

Las propuestas presentadas seguirán **la siguiente estructura**:

1. Información sobre la **entidad que realiza la oferta**:
 - NIF y razón social. La entidad deberá tener capacidad para desarrollar las tareas solicitadas.
 - Fecha en la que se realiza la oferta y firma por la empresa.
 - **Solvencia técnica y profesional**, indicando la relación de los principales servicios o trabajos similares realizados en los últimos tres años, incluyendo importe, fechas y destinatario, tal y como se describe en el apartado 3.3.

2. **Memoria técnica**, en la que se detallen el enfoque general, los métodos de trabajo y las herramientas previstas para el desarrollo del proyecto. Esta memoria deberá:
 - Incluir una **propuesta técnica detallada**, que contemple la aproximación metodológica planteada para alcanzar los objetivos del proyecto.
 - Reflejar de forma clara y estructurada la **trazabilidad respecto a los paquetes de trabajo, actividades y entregables** establecidos.
3. **Presupuesto detallado**, que deberá respetar el límite máximo establecido en el presente pliego, desglosando los costes por paquete de trabajo o fase, e incluyendo todos los conceptos necesarios para la ejecución del proyecto.
4. **Plan de trabajo, cronograma y asignación de recursos**. Este apartado de la propuesta ofrecerá los siguientes contenidos:
 - **Plazo de ejecución propuesto**, que deberá ajustarse a los límites establecidos en el presente pliego.
 - **Plan de trabajo**, que incluirá el cronograma detallado de las actividades, la secuencia de los paquetes de trabajo y la asignación de recursos humanos y materiales.
 - **Composición del equipo de trabajo**, describiendo la experiencia, perfil profesional y funciones del personal que se asignará al proyecto.

6.3. Criterios de adjudicación

La adjudicación del presente contrato se realizará mediante la aplicación de los siguientes criterios de adjudicación, cuantitativos y cualitativos, que alcanzarán un máximo de 100 puntos:

A) Oferta técnica: Hasta 50 puntos.

Cada licitador deberá presentar su **oferta** que deberá incluir todos los elementos necesarios para valorar la calidad de esta, conforme a la estructura propuesta en el apartado anterior.

B) Precio: Hasta 20 puntos.

Se asignarán puntos en función de la baja en el precio del contrato, conforme la fórmula siguiente:

$$\text{Puntuación} = \frac{20 \cdot b^2}{20 + b^2}$$

b= porcentaje de baja, en número. (Ej.: para conocer la puntuación de un 5,15% de baja, colocar el número 5,15 en lugar de “b”).

c) Plan de trabajo, cronograma y asignación de recursos: Hasta 30 puntos.

Cada licitador deberá disponer de los medios necesarios y de personal técnico suficiente y cualificado, para garantizar el desarrollo de los trabajos detallados en el presente pliego.

En este apartado de la propuesta se valorará la disponibilidad de medios aplicados a la ejecución del trabajo, en especial la polivalencia y adecuación del equipo humano. Se incluirán los CV resumidos de las personas integrantes del equipo.

Se valorará la experiencia acumulada en proyectos de naturaleza similar, particularmente en el ámbito de la industria musical, la gestión de proyectos financiados con fondos públicos o la participación en iniciativas de apoyo a la transformación digital de organizaciones.

6.4. Adjudicación

La adjudicación del contrato se acordará por el Comité de Dirección de AIE con arreglo al informe de propuesta de adjudicación a la oferta con mejor valoración, realizado por el equipo designado por parte de AIE, utilizando los criterios de adjudicación, tanto cuantitativos como cualitativos.

En el caso de que no se hayan presentado ofertas o las presentadas sean inadecuadas, irregulares o inaceptables el procedimiento se declarará desierto. Las ofertas que no cumplan con lo indicado en los pliegos se considerarán no aptas. En estos casos no habrá derecho a indemnización alguna.

6.5. Formalización

El contrato se formalizará por escrito en los 30 días siguientes al acuerdo de adjudicación. Tanto los adjudicatarios, como las personas de AIE que participen en el proceso de licitación han de firmar una declaración de ausencia de conflicto de interés (DACI), con carácter previo a la firma del contrato.

A efectos de la formalización del contrato, AIE podrá requerir aquella documentación que estime necesaria al adjudicatario, sin que se proceda a la formalización hasta la completa remisión de la misma.

Adicionalmente, los adjudicatarios deberán cumplir con las obligaciones que se derivan del Reglamento del Mecanismo de Recuperación y Resiliencia (Reglamento (UE) 2021/241) respecto a la lucha contra el fraude, en materia de comunicación y el cumplimiento del principio DNSH.

La formalización del contrato implicará la aceptación inequívoca y total de estos pliegos, así como de los compromisos asumidos en la propuesta, siendo ley entre las partes

7. MODIFICACIONES

En caso de que se produzcan eventualidades que afecten a la planificación, organización del proyecto o composición del equipo de trabajo, será AIE quien autorice, en su caso, las modificaciones más adecuadas, de acuerdo con lo establecido en el artículo 37 de la Orden de bases de las ayudas para Espacios de Datos publicada por el Ministerio de Cultura. En ningún caso el adjudicatario podrá introducir cambios sin la autorización expresa y previa de AIE.

8. OBLIGACIONES FORMALES DEL ADJUDICATARIO

El adjudicatario se compromete a cumplir las siguientes obligaciones formales relativas al cumplimiento de los principios de gestión del PRTR:

- Con **carácter previo a la firma del contrato**, una vez comunicada la adjudicación, el contratista y, en su caso, los subcontratistas, deberán firmar una declaración de ausencia de conflicto de interés (DACI), siguiendo el modelo que se adjunta como anexo I a este pliego.
- También con **anterioridad a la firma del contrato**, una vez comunicada la adjudicación, la empresa contratista y, en su caso, las subcontratistas, deberán firmar una declaración de compromiso en relación con la ejecución de actuaciones del Plan de recuperación, transformación y resiliencia (PRTR), siguiendo el modelo que se adjunta en el Anexo II a este pliego.
- **Comunicación y Visibilidad de la Financiación:** la empresa adjudicataria se compromete a mencionar la financiación de la Unión Europea y dar visibilidad mediante el emblema de la Unión y la declaración de financiación adecuada: **"financiado por la Unión Europea – NextGenerationEU"**. Esto se aplicará en todos los documentos y comunicaciones relativos al objeto del contrato.
- **Uso de Logotipos:** Todos los documentos y entregables deberán incluir en un lugar visible (preferiblemente portada o encabezado) una tira de logos que contenga:
 - El emblema de la Unión Europea, acompañado de la declaración "financiado por la Unión Europea – Next generationEU"
 - El logotipo del Ministerio para la Transformación Digital y de la Función Pública.
 - El logotipo del Plan de Recuperación, Transformación y Resiliencia.
 - El logotipo de AIE.

- **Presentación de Documentación:** La empresa adjudicataria deberá presentar toda la documentación que le requiera AIE para comprobar el correcto cumplimiento del objeto del presente contrato.

9. CONDICIONES DE FACTURACIÓN Y PAGO

Todas las facturas deberán emitirse cumpliendo con los requerimientos fiscales y mercantiles impuestos por la legislación vigente y se presentarán en el lugar, moneda, condiciones de pago y plazos que se determinen en el contrato que las partes suscriban para la formalización de la adjudicación de los trabajos objeto de contratación.

En particular:

- Las facturas deberán cumplir con lo expuesto en el Real Decreto 1619/2012, de 30 de noviembre, por el que se aprueba el reglamento por el que se regulan las obligaciones de facturación.
- Las facturas deberán tener fecha de emisión dentro del plazo de ejecución del proyecto.
- Cuando se trate de bienes, se deberá identificar claramente el producto (marca, modelo, número de serie), así como el coste unitario del mismo.
- Cuando se trate de servicios, se indicará el contrato asociado, en su caso, las tareas realizadas, debidamente detalladas y en consonancia con las actividades del proyecto, los costes/hora por perfil facturados y el número de horas correspondiente. En caso de ser necesario, se podrá admitir un anexo a la factura debidamente firmado por el proveedor emisor de la misma.

La facturación se realizará con carácter mensual, en función del avance efectivo de los trabajos ejecutados, conforme a la planificación establecida para las distintas actividades y entregables del proyecto. **Cada factura deberá hacer referencia expresa a los conceptos descritos en el apartado 4 del presente pliego**, con el fin de garantizar la adecuada trazabilidad entre los hitos facturados y los trabajos realizados. Asimismo, junto con cada factura, **el adjudicatario deberá aportar un documento justificativo en el que se detallen los trabajos realizados durante el periodo correspondiente**, vinculándolos con las actividades y entregables a los que se refiera la factura. Esta documentación será condición necesaria para la validación y aceptación del pago por parte de AIE.

Por otro lado, cada factura deberá incluir la mención siguiente:

"Este proyecto será financiado por la Unión Europea-Next Generation EU, al amparo de la convocatoria de ayudas Espacios de Datos Sectoriales 2023 para la creación de demostradores y casos de uso de espacios de compartición de datos, y de conformidad con lo dispuesto en el artículo 34 de la Orden TDF/1461/2023, de 29 de diciembre, por la

que se aprueban las bases reguladoras y se efectúa la convocatoria para la concesión de ayudas, en el ámbito de la digitalización, para la transformación digital de los sectores productivos estratégicos mediante la creación de demostradores y casos de uso de Espacios de Compartición de Datos, en el marco del Plan de Recuperación, Transformación y Resiliencia-Next Generation EU (Programa Espacios de Datos Sectoriales), modificada por la Orden TDF/294/2024, de 4 de abril".

Los pagos se realizarán por AIE en un plazo de 30 días contados a partir del día siguiente de la fecha de la factura, mediante transferencia bancaria al número de cuenta que el adjudicatario indique.

10. SEGURIDAD, CONFIDENCIALIDAD Y PROTECCIÓN DE DATOS

AIE proveerá a la empresa adjudicataria de toda la información necesaria para realizar de los trabajos asignados. A estos efectos, se adjuntará un contrato de encargo de tratamiento de datos personales, en virtud del cual la empresa adjudicataria pasará a tener la condición de encargado de tratamiento de datos de carácter personal.

El adjudicatario queda expresamente obligado a mantener absoluta confidencialidad y reserva sobre cualquier dato que pudiera conocer con ocasión del cumplimiento del contrato, especialmente los de carácter personal, que no podrá copiar ni utilizar con una finalidad distinta a la que figura en este documento, ni tampoco ceder a otros ni siquiera a efectos de conservación.

Este deber subsistirá, aunque se extinga el contrato, hasta que la información llegue a ser de dominio público o que, por otras causas, pierda su consideración de confidencialidad, sin que medie incumplimiento del deber de confidencialidad por cualquiera de las partes contratantes.

La empresa adjudicataria se obliga al cumplimiento de la normativa vigente en materia de protección de datos, Reglamento (UE) 2016/679 del Parlamento Europeo y del Consejo de 27 de abril de 2016 (GDPR) y Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales.

Asimismo, se obliga en los términos previstos en la Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales, a garantizar total confidencialidad y seguridad de los datos a los que tenga acceso para su tratamiento.

Los trabajos objeto de este contrato deberán ser convenientemente documentados, para lo cual el adjudicatario se compromete a generar toda la documentación aplicable.

11. PROPIEDAD DE LOS TRABAJOS REALIZADOS

Los trabajos contratados, en su totalidad o en cualquiera de sus fases, serán propiedad de AIE y, en consecuencia, podrán solicitar en cualquier momento las entregas de los documentos o materiales que la integren, con todos sus antecedentes, datos o procedimientos en cualquier tipo de formato editable.

En consecuencia, los trabajos que sean objeto de propiedad intelectual se entenderán expresamente cedidos en exclusiva por el adjudicatario que, en su caso, cuidará de que se cedan por terceros en exclusiva al AIE.

12. CONTACTO

Para cualquier duda en relación con estos pliegos y el procedimiento de contratación, pueden dirigir un correo electrónico a secregen@aie.es

En Madrid, a 26 de junio de 2025

ANEXO I. DECLARACIÓN DE AUSENCIA DE CONFLICTO DE INTERÉS (DACI)

Orden TDF/1461/2023, de 29 de diciembre, por la que se aprueban las bases reguladoras y se efectúa la convocatoria para la concesión de ayudas, en el ámbito de la digitalización, para la transformación digital de los sectores productivos estratégicos mediante la creación de demostradores y casos de uso de Espacios de Compartición de Datos, en el marco del Plan de Recuperación, Transformación y Resiliencia-Next Generation EU (Programa Espacios de Datos Sectoriales, C12, I1).

DECLARACIÓN DE AUSENCIA DE CONFLICTO DE INTERESES CONTRATISTA (DACI)

D./Dña. _____ con NIF _____ y domicilio fiscal en _____, con capacidad legal necesaria en virtud de _____, por haber resultado adjudicatario del contrato _____, licitado y tramitado en el marco de la convocatoria de para la concesión de ayudas, en el ámbito de la digitalización, para la transformación digital de los sectores productivos estratégicos mediante la creación de demostradores y casos de uso de Espacios de Compartición de Datos, en el marco del Plan de Recuperación, Transformación y Resiliencia-Next Generation EU (Programa Espacios de Datos Sectoriales).

DECLARA

Primero. Estar informado/s de lo siguiente:

1. Que el artículo 61.3 «*Conflicto de intereses*», del Reglamento (UE, Euratom) 2018/1046 del Parlamento Europeo y del Consejo, de 18 de julio (Reglamento financiero de la UE) establece que «*existirá conflicto de intereses cuando el ejercicio imparcial y objetivo de las funciones se vea comprometido por razones familiares, afectivas, de afinidad política o nacional, de interés económico o por cualquier motivo directo o indirecto de interés personal.*»

2. Que el artículo 64 «*Lucha contra la corrupción y prevención de los conflictos de intereses*» de la Ley 9/2017, de 8 de noviembre, de Contratos del Sector Público, tiene el fin de evitar cualquier distorsión de la competencia y garantizar la transparencia en el procedimiento y asegurar la igualdad de trato a todos los candidatos y licitadores.

3. Que el artículo 23 «*Abstención*», de la Ley 40/2015, de 1 octubre, de Régimen Jurídico del Sector Público, establece que deberán abstenerse de intervenir en el procedimiento «*las autoridades y el personal al servicio de las Administraciones en quienes se den algunas de las circunstancias señaladas en el apartado siguiente*», siendo éstas:

a) Tener interés personal en el asunto de que se trate o en otro en cuya resolución pudiera influir la de aquél; ser administrador de sociedad o entidad interesada, o tener cuestión litigiosa pendiente con algún interesado.

b) Tener un vínculo matrimonial o situación de hecho asimilable y el parentesco de consanguinidad dentro del cuarto grado o de afinidad dentro del segundo, con cualquiera de los interesados, con los administradores de entidades o sociedades interesadas y también con los asesores, representantes legales o mandatarios que intervengan en el procedimiento, así como

compartir despacho profesional o estar asociado con éstos para el asesoramiento, la representación o el mandato.

c) Tener amistad íntima o enemistad manifiesta con alguna de las personas mencionadas en el apartado anterior.

d) Haber intervenido como perito o como testigo en el procedimiento de que se trate.

e) Tener relación de servicio con persona natural o jurídica interesada directamente en el asunto, o haberle prestado en los dos últimos años servicios profesionales de cualquier tipo y en cualquier circunstancia o lugar.

Segundo. Que no se encuentra/n incurso/s en ninguna situación que pueda calificarse de conflicto de intereses de las indicadas en el artículo 61.3 del Reglamento Financiero de la UE y que no concurre en su/s persona/s ninguna causa de abstención del artículo 23.2 de la Ley 40/2015, de 1 de octubre, de Régimen Jurídico del Sector Público que pueda afectar al procedimiento de licitación/concesión.

Tercero. Que se compromete/n a poner en conocimiento del órgano de contratación/comisión de evaluación, sin dilación, cualquier situación de conflicto de intereses o causa de abstención que dé o pudiera dar lugar a dicho escenario.

Cuarto. Conozco que, una declaración de ausencia de conflicto de intereses que se demuestre que sea falsa, acarreará las consecuencias disciplinarias/administrativas/judiciales que establezca la normativa de aplicación.

(Fecha y firma, nombre completo y DNI del contratista/subcontratista)

ANEXO II. DECLARACIÓN DE COMPROMISO EN RELACIÓN CON LA EJECUCIÓN DE ACTUACIONES DEL PRTR

Orden TDF/1461/2023, de 29 de diciembre, por la que se aprueban las bases reguladoras y se efectúa la convocatoria para la concesión de ayudas, en el ámbito de la digitalización, para la transformación digital de los sectores productivos estratégicos mediante la creación de demostradores y casos de uso de Espacios de Compartición de Datos, en el marco del Plan de Recuperación, Transformación y Resiliencia-Next Generation EU (Programa Espacios de Datos Sectoriales).

DECLARACIÓN DE COMPROMISO EN RELACIÓN CON LA EJECUCIÓN DE ACTUACIONES DEL PLAN DE RECUPERACIÓN, TRANSFORMACIÓN Y RESILIENCIA (PRTR).

D./Dña. _____ en
representación _____ de _____ con
NIF _____, con capacidad legal necesaria en virtud
de _____, que participa como contratista/subcontratista, en el desarrollo
de actuaciones necesarias para la consecución de los objetivos del Componente 12 “Política Industrial España 2030”, Inversión 1 “Espacios de datos Sectoriales”, manifiesta el compromiso de la entidad que representa con los estándares más exigentes en relación con el cumplimiento de las normas jurídicas, éticas y morales, adoptando las medidas necesarias para prevenir y detectar el fraude, la corrupción y los conflictos de interés, comunicando en su caso a las autoridades que proceda los incumplimientos observados.

Adicionalmente, atendiendo al contenido del PRTR, se compromete a respetar los principios transversales del PRTR, entre ellos el de economía circular y evitar impactos negativos significativos en el medio ambiente («DNSH» por sus siglas en inglés «do no significant harm») en la ejecución de las actuaciones llevadas a cabo en el marco de dicho Plan, y manifiesta que no incurre en doble financiación y que, en su caso, no le consta riesgo de incompatibilidad con el régimen de ayudas de Estado.

_____, XX de _____ de 202X

Firma del Contratista/Subcontratista

ANEXO III. ANÁLISIS DE CONTEXTO. ARQUITECTURA DEL ESPACIO DE DATOS

Arquitectura del Espacio de Datos

2025-04-30

Índice

1. Introducción	4
1.A. Iniciativas Relacionadas con Espacio de Datos y Selección del Marco IDSA.....	4
2. Principios de la Arquitectura	7
2.A. Soberanía y Gobernanza del Dato	7
2.B. Capa Confiable de Intercambio de Datos	7
2.C. Almacenamiento de Datos.....	7
2.D. Privacidad y Seguridad.....	8
2.E. Monitoreo y Auditoría.....	8
2.F. Transformación de Datos.....	8
2.G. Accesibilidad de la Información	9
3. Arquitecturas de Referencia.....	9
3.A. Arquitectura IDSA RAM 4.0	9
Componentes principales del IDS-RAM.....	10
Implementación y Recursos	11
3.B. Arquitectura de Otros Espacios de Datos Europeos	11
4. Arquitectura del Espacio de Datos	13
4.A. Selección de Proveedores Tecnológicos.	15
Justificación de componentes Eclipse Dataspace y uso de AWS.....	15
¿Por qué AWS y no otras plataformas como Gaia-X?	15
4.B. Detalle de Arquitectura	17
Dominio del Participante	18
Dominio del Espacio de Datos.....	21
4.C. Integración y Despliegue de la Arquitectura	29
Integración con el Espacio de Datos.....	29
Integración de los Casos de Uso	31
5. Procesos.....	32
5.A. Gestión de Participantes	32

5.B. Despliegue de Componentes y Operaciones en el Espacio de Datos	33
5.C. Gestión de Aplicaciones y Casos de Uso	33
6. Aspectos no Funcionales	34
6.A. Seguridad y Privacidad	34
6.B. Rendimiento y Escalabilidad	35
Common Dataspace	35
Participant Dataspace.....	35
6.C. Despliegue e Integración.....	36
6.D. Mantenimiento y Soporte.....	37
7. Pruebas de Concepto para la validación de la arquitectura propuesta	37
8. Beneficios de la Arquitectura en IDS	37
9. Roadmap para la Implementación Práctica de EDME.....	38
10. Conclusión	40

1. Introducción

El **principal objetivo del proyecto EDME** consiste en el desarrollo de un Espacio de Datos (ED) para la música española donde los principales actores de la industria puedan intercambiar datos de una manera segura, transparente y soberana. De este modo se pretende explotar en su máxima expresión la economía del dato rompiendo con los silos de información existentes actualmente en esta y otras industrias. Asimismo, se desarrollarán diferentes casos de uso representativos con el fin de evaluar el impacto que pueden generar estos Espacios de Datos.

El espacio de datos internacional, según los principios de la **International Data Spaces Association (IDSA)**, permite la creación de un entorno de intercambio de datos que promueve la **soberanía y confianza** en los datos, respetando la privacidad y seguridad de los mismos. La implementación de un espacio de datos a nivel global facilita la colaboración entre organizaciones sin comprometer el control de los datos. Cada participante mantiene el control absoluto sobre sus datos, decidiendo cuándo, cómo y con quién compartirlos. Este modelo también garantiza el cumplimiento de las regulaciones globales sobre protección de datos, como el **Reglamento General de Protección de Datos (GDPR)** de la Unión Europea.

El enfoque de IDSA se basa en una **arquitectura descentralizada** donde los datos no se centralizan, sino que permanecen bajo el control de los proveedores de datos, quienes pueden compartir solo los datos que elijan, bajo las condiciones que establezcan. Además, los procesos de intercambio de datos se llevan a cabo en un entorno **seguro y transparente**, con mecanismos de autenticación, autorización y monitoreo que garantizan la privacidad y la integridad de los datos.

En este documento se describe cómo implementar una arquitectura de espacio de datos basado en los principios de IDSA en la nube de **AWS**. Asimismo, se introducen diferentes términos asociados al concepto de ED y se describe **su implementación práctica y operativa**. La arquitectura propuesta se estructura para permitir un intercambio eficiente de datos entre las organizaciones, mientras se mantiene el control total de los datos por parte de los propietarios. Además, se destacan los beneficios de utilizar AWS como plataforma, aprovechando sus servicios de almacenamiento, procesamiento, seguridad y gestión de identidades para crear un entorno flexible y escalable para los participantes.

1.A. Iniciativas Relacionadas con Espacio de Datos y Selección del Marco IDSA

Durante los últimos años han surgido **múltiples iniciativas entorno al ED** destacando:

- *Dataspace Support Center (DSSC)*: se trata de un proyecto europeo de 3 años financiado dentro del programa Digital Europe. Este proyecto explora las necesidades existentes en los actuales programas de ED, define requisitos comunes y establece buenas prácticas con el fin de acelerar el proceso de formación de ED soberanos.

- *International Data Spaces Association (IDSA)*: su visión es crear un entorno de confianza para promover el intercambio de datos a través de un ED federado, internacional y globalmente certificado. El principal objetivo es crear un estándar global para los ED, así como promover las tecnologías y modelos de negocio que permitan explotar la economía del dato.
- *Gaia-X*: Su objetivo es crear una infraestructura de datos abierta federada basada en los valores europeos de soberanía del dato y la nube. Su misión es diseñar e implementar una arquitectura de compartición de datos que esté basada en estándares comunes para el intercambio de los datos, las mejores prácticas, herramientas y mecanismos de gobernanza
- *DSBA*: Este documento ha sido publicado por la *Data Spaces Business Alliance (DSBA)*, una coalición de cuatro asociaciones europeas que promueven el desarrollo de ED. El documento define un marco tecnológico de referencia para la creación de ED en base a la convergencia técnica de los modelos y arquitecturas existentes. Así describe los principales pilares tecnológicos que habrán de disponibilizar los ED, tales como la interoperabilidad de datos, la soberanía y la confianza y la creación de valor a partir del dato.
- *FIWARE*: La fundación Fiware ha desarrollado un marco de componentes de plataforma de código abierto para acelerar el desarrollo de soluciones inteligentes. Este proporciona muchos de los *building blocks* requeridos para completar el ecosistema de compartición de datos.
- *Simpl*: Se trata de una plataforma *middleware* inteligente promovida por la Comisión Europea dentro del programa Digital Europe. Su principal objetivo es proporcionar a los operadores de espacio de datos con un denominador técnico común que asegure niveles adecuados de confianza, seguridad, facilidad en el acceso, adaptabilidad e interoperabilidad.

En este trabajo **se ha optado por seguir las directrices marcadas por la IDSA por diferentes razones:**

- Es uno de los principales estándares internacionales para el desarrollo de ED, contando con un amplio abanico de documentación, estándares, certificaciones o componentes de código abierto que apoyan el desarrollo de nuevas iniciativas de ED.
- Entre los miembros de la asociación IDSA cuenta con las principales universidades y centros tecnológicos dentro del área de la tecnología de datos, así como grandes corporaciones internacionales como Google, Volkswagen o Huawei
- Buen alineamiento entre los diferentes componentes recomendados por IDSA y el ecosistema tecnológico AWS en el que se implementará EDME.
- Buena concordancia entre los principios promovidos por IDSA y los que guiarán el desarrollo de EDME, que podemos resumir en los siguientes puntos:

- “No reinventar la la rueda”, empleando tecnologías que ya estén disponibles y testadas
- Integración de ED en sistemas ya existentes, en la medida de lo posible
- Integración y uso de estándares disponibles, buscando el alineamiento entre las especificaciones y los estándares técnicos
- Fácil uso con énfasis en su portabilidad y replicabilidad

2. Principios de la Arquitectura

2.A. Soberanía y Gobernanza del Dato

El principio fundamental detrás de la **soberanía de los datos** es que cada organización que participa en el espacio de datos mantiene el control total sobre sus datos. Esto implica que pueden decidir con quién compartir los datos, qué tipo de acceso permitir y bajo qué condiciones. La soberanía no solo se refiere al control sobre los datos en sí, sino también a las **políticas de acceso** que se aplican a esos datos. Las organizaciones pueden configurar y definir qué reglas se deben seguir para que otros puedan consumir sus datos, asegurando que siempre se respeten sus condiciones.

La **gobernanza** en un espacio de datos IDSA es un aspecto esencial para asegurar que el intercambio de datos se realice bajo las normativas adecuadas. Un marco de gobernanza bien definido asegura que todos los participantes sigan un conjunto de reglas que definan cómo se compartirá, procesará y usará la información. Esto incluye la **gestión de permisos**, la **verificación de la autenticidad de los participantes** y el **cumplimiento de normativas locales e internacionales**. El marco de gobernanza debe ser transparente, auditable y fácilmente adaptable a nuevas regulaciones.

2.B. Capa Confiable de Intercambio de Datos

El **Data Space Connector (DSC)** es el componente clave que permite la integración entre los diferentes participantes del espacio de datos. El DSC actúa como una capa de intermediación, facilitando el **intercambio seguro** de datos entre proveedores y consumidores. Funciona como un puente que asegura que los datos no solo se transmiten de manera segura, sino que se intercambian de acuerdo con las reglas y políticas establecidas en el marco de gobernanza.

El DSC garantiza que los datos nunca abandonen el control del proveedor de datos, sino que solo se compartan bajo las condiciones establecidas. Además, la autenticación y autorización dentro del DSC se manejan mediante un **Identity Provider (IdP)** y protocolos seguros como OAuth 2.0. Esto asegura que solo las entidades autenticadas y autorizadas tengan acceso a los datos, lo que aumenta la confianza en el espacio de datos.

2.C. Almacenamiento de Datos

El almacenamiento descentralizado es un principio clave en la arquitectura de IDSA. En lugar de almacenar los datos de manera centralizada en un único repositorio, los datos permanecen en el **dominio del participante** y solo se intercambian de manera controlada cuando se permite. Los participantes del espacio de datos gestionan su propio almacenamiento, lo que les permite mantener el control total sobre sus datos.

En la fase inicial de la implementación de EDME, se propone el uso de **Amazon S3** para proporcionar almacenamiento a los participantes. Cada participante podrá almacenar sus datos en un **espacio seguro en S3**, con **políticas de acceso definidas** a través de **AWS IAM** para garantizar que solo los participantes autorizados tengan acceso. Este enfoque reduce las barreras de entrada para los nuevos participantes, ya que no tienen que preocuparse por implementar infraestructura de almacenamiento compleja.

A medida que EDME crezca, los participantes podrán gestionar sus propios entornos de almacenamiento descentralizados, manteniendo así la soberanía sobre los datos, pero beneficiándose de una arquitectura flexible y escalable.

2.D. Privacidad y Seguridad

La **seguridad** y la **privacidad** son aspectos fundamentales en cualquier arquitectura de intercambio de datos. Todos los datos, tanto en tránsito como en reposo, deben estar **cifrados** utilizando protocolos seguros. **AWS KMS** es una herramienta esencial para gestionar el cifrado de los datos en reposo, mientras que los **protocolos TLS** garantizan la seguridad de los datos en tránsito entre los diferentes nodos del espacio de datos.

Además, el control de acceso se gestiona mediante **AWS IAM**, que permite establecer **políticas de acceso granular** a los recursos. Esto asegura que solo las personas o aplicaciones autorizadas puedan acceder a los datos, lo que proporciona una capa adicional de seguridad.

2.E. Monitoreo y Auditoría

El monitoreo y la auditoría son cruciales para garantizar que el intercambio de datos se realiza de acuerdo con las reglas establecidas y para detectar cualquier actividad sospechosa o no autorizada. **AWS CloudTrail** se utiliza para registrar todas las actividades realizadas dentro del espacio de datos sobre el almacenamiento primario (s3), proporcionando un registro completo de las interacciones. Este registro es esencial para **auditorías** y para mantener la transparencia en todo el proceso de intercambio de datos.

Amazon CloudWatch se utiliza para monitorear en tiempo real el **flujo de datos** y las **actividades del sistema**, alertando a los administradores sobre posibles problemas o incidencias. Este enfoque asegura que todos los eventos sean rastreables y auditables, lo que facilita el cumplimiento de las normativas de seguridad y privacidad.

2.F. Transformación de Datos

La **transformación de datos** se refiere a la capacidad de manipular y transformar los datos dentro del espacio de datos para satisfacer las necesidades de los consumidores. En la arquitectura propuesta, se incluye una **tienda de aplicaciones** de datos que permite a los

participantes distribuir aplicaciones que realicen operaciones sobre los datos, como análisis o transformación. Aunque en las primeras fases no se planea un desarrollo profundo en este área, se propone el uso de **contenedores OCI** para empaquetar estas aplicaciones.

Amazon ECR puede ser utilizado como registro de contenedores para almacenar y gestionar estos contenedores, permitiendo que las aplicaciones sean fácilmente distribuidas y ejecutadas por los participantes del espacio de datos. A largo plazo, la implementación de **contenedores y microservicios** puede ofrecer un enfoque más modular y escalable para la transformación de datos.

2.G. Accesibilidad de la Información

Un aspecto crítico del espacio de datos es la **accesibilidad** de la información. Para facilitar la navegación de los usuarios dentro del espacio de datos, se propone la creación de un **Data Marketplace**, un servicio web con un **front-end intuitivo** que permita a los participantes gestionar fácilmente los datasets, las políticas de acceso y las aplicaciones disponibles.

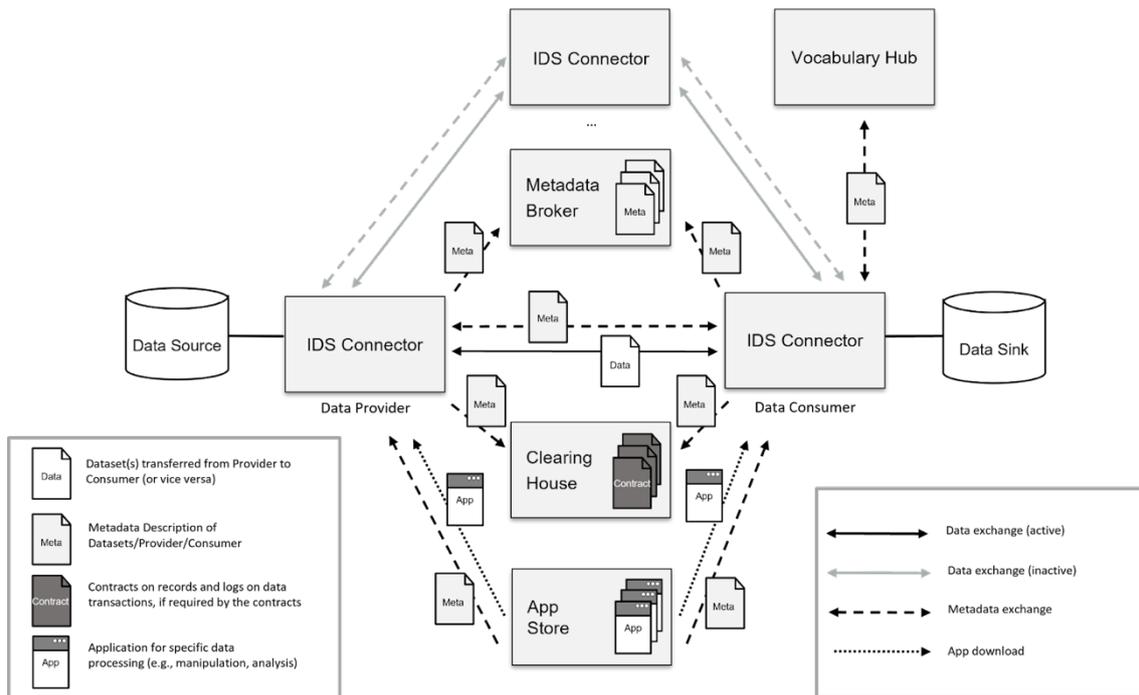
El **Data Marketplace** actúa como un punto central donde los usuarios pueden buscar, acceder y consumir datos, además de visualizar las políticas de privacidad y las reglas de acceso asociadas. Eventualmente, los participantes pueden optar por desarrollar sus propios frontales web, pero la implementación inicial del marketplace ayudará a reducir las barreras de entrada y facilitará la adopción del espacio de datos.

3. Arquitecturas de Referencia

El diseño de la arquitectura del espacio de datos EDME sigue las directrices de la arquitectura de referencia IDS-RAM 4.0, propuesta por IDSA, si bien también se han tenido en cuenta otras iniciativas de espacios de datos a nivel europeo y nacional, como comentaremos a continuación.

3.A. Arquitectura IDSA RAM 4.0

El modelo de arquitectura de referencia de la International Data Spaces Association (IDS-RAM 4.0) define un conjunto de componentes fundamentales que permiten el intercambio seguro, confiable y soberano de datos entre organizaciones. Estos componentes trabajan en conjunto para garantizar la autenticación de los participantes, la protección de los datos, la aplicación de políticas de uso y la interoperabilidad técnica y semántica dentro del ecosistema IDS.



Componentes principales del IDS-RAM

Dentro de la arquitectura IDS RAM podemos distinguir los siguientes componentes, no todos ellos obligatorios para la implementación mínima de un espacio de datos. Las piezas imprescindibles incluyen:

1. **IDS Connector** ♦ *Obligatorio para todos los participantes*
Es el núcleo del ecosistema IDS. Sin él, no se puede intercambiar datos ni aplicar políticas de uso.
2. **Identity Provider (IdP)** ♦ *Obligatorio a nivel de sistema*
Todos los conectores deben autenticar su identidad mediante un proveedor de identidad certificado por IDSA.
3. **Policy Enforcement Mechanism (en el conector)** ♦ *Obligatorio dentro del conector*
Debe existir un mecanismo para aplicar y respetar las políticas de uso de datos acordadas.

Los siguientes componentes no son estrictamente necesarios, pero recomendables para enriquecer la funcionalidad, la trazabilidad o la interoperabilidad,

1. **Metadata Broker** ● *Opcional pero altamente recomendado*
Facilita el descubrimiento de conectores y ofertas de datos. No es obligatorio si los conectores se conocen directamente.

2. **Clearing** **House** ● Opcional
Aporta trazabilidad y auditoría, útil en entornos con requisitos regulatorios o de confianza reforzada.
3. **App** **Store** ● Opcional
Permite ejecutar apps certificadas sobre los datos. Útil si se requiere procesamiento dentro del conector.
4. **Vocabulary** **Provider** / **Ontology** **Hub** ● Opcional
Soporta interoperabilidad semántica, muy útil en contextos intersectoriales o multilingües.
5. **Certification Service** ● Opcional (pero necesario para participantes certificados)
Aunque no todos los participantes necesitan certificarse, es obligatorio para quienes operan bajo requisitos formales de seguridad y calidad.

Uno de los componentes clave del esquema IDS-RAM es el Conector IDS. Este actúa como una puerta de enlace que permite a las organizaciones conectar sus sistemas existentes al ecosistema IDS, facilitando el intercambio de datos de manera segura y controlada. El conector gestiona aspectos como la autenticación, autorización, gestión de metadatos y aplicación de políticas de uso de datos.

Implementación y Recursos

El IDS-RAM es un modelo de código abierto y su documentación y especificaciones están disponibles para su consulta y contribución en plataformas como GitHub. Esto permite a las organizaciones y desarrolladores acceder al modelo, adaptarlo a sus necesidades específicas y participar en su evolución continua.

3.B. Arquitectura de Otros Espacios de Datos Europeos

También se han estudiado y analizado otros espacios de datos europeos y nacionales con el fin de evaluar no solamente la arquitectura, sino también la operativa y los modelos de negocio que siguen estas iniciativas. En concreto cabe destacar las siguientes:

- **Catena-X:** es una iniciativa que busca crear un ecosistema de datos federado y seguro para la **industria automotriz**, facilitando la colaboración y el intercambio de información entre empresas a lo largo de toda la cadena de valor. Dentro de este espacio de datos cabe destacar:
 - **Servicios de Habilitación (Enablement Services):** Estos servicios pueden ser desplegados de forma autónoma por cada participante o gestionados por un proveedor de servicios de habilitación. Su función principal es facilitar la provisión y el consumo de datos dentro del espacio de datos de Catena-X.

- **Aplicaciones Empresariales (Business Applications):** Estas aplicaciones abordan problemas empresariales específicos, como la gestión de la demanda y la capacidad, o la economía circular. Pueden variar desde soluciones empresariales de gran escala hasta soluciones especializadas para pequeñas y medianas empresas (vía Application Stores).
- **Copernicus (CDSE):** es una infraestructura avanzada que facilita el acceso y procesamiento de datos de observación terrestre provenientes de las misiones Sentinel y otras fuentes. Su objetivo es proporcionar a usuarios de diversos sectores una plataforma integrada para explorar, analizar y utilizar información geoespacial de manera eficiente.
 - **Acceso a Datos:** El CDSE ofrece acceso libre e inmediato a una amplia gama de datos de observación de la Tierra, incluyendo imágenes satelitales de las misiones Sentinel y otras contribuciones. Estos datos están disponibles para su visualización y descarga a través del [Copernicus Browser](#), una herramienta intuitiva que permite a los usuarios explorar y seleccionar la información de interés.
 - **Herramientas de Procesamiento:** Además del acceso a datos, el ecosistema proporciona un entorno robusto para el análisis y procesamiento de información geoespacial. Incluye herramientas y servicios que permiten a los usuarios extraer información valiosa, realizar modelados y generar productos personalizados basados en los datos disponibles.
 - **Ecosistema Colaborativo:** El CDSE actúa como un hub que integra herramientas y recursos destinados a desbloquear el potencial completo de los datos de observación terrestre. Fomenta la colaboración entre desarrolladores, investigadores y empresas, facilitando la creación y compartición de aplicaciones y servicios que aprovechan la información proporcionada por Copernicus.
- Proyecto Horizon Europe **Open Music Europe:** su objetivo es crear el prototipo de un Observatorio Europeo de la Música moderno, abierto, transparente y altamente automatizado. Pretende garantizar que todos los actores representativos de la música a nivel internacional y de la UE, así como las organizaciones musicales nacionales y regionales, puedan aprovechar al máximo los datos, software y algoritmos abiertos dispuestos por el proyecto. Se trata de una acción preparatoria para la construcción de un Espacio Europeo para la Música.
- **Plataforma Inteligente de Destinos (SEGITTUR)** y Espacio de Datos Europeo para el Turismo (DATES y Tourism Data Space). Se trata de uno de los ED más maduros a nivel nacional y por tanto ha sido relevante su análisis tecnológico para extraer ideas que nos pudieran ayudar a diseñar la futura arquitectura de EDME. La plataforma pretende integrar, relacionar y combinar datos públicos y privados para generar más inteligencia competitiva en los ecosistemas de los destinos españoles, activando la innovación

continua, interconectando y atendiendo las necesidades de todos los agentes de la cadena de valor turístico: turista, destinos y empresas. Asimismo, ha desplegado pruebas de concepto para casi una decena de casos de uso. Alineado con esta plataforma para el turismo, también se han evaluado dos proyectos europeos, DATES y Tourism Data Space, cuyos enfoques están dirigidos a la definición del marco para el desarrollo de un espacio de datos de turismo europeo y a la participación activa de los distintos agentes que sirva de base para el diseño del espacio de datos de turismo europeo.

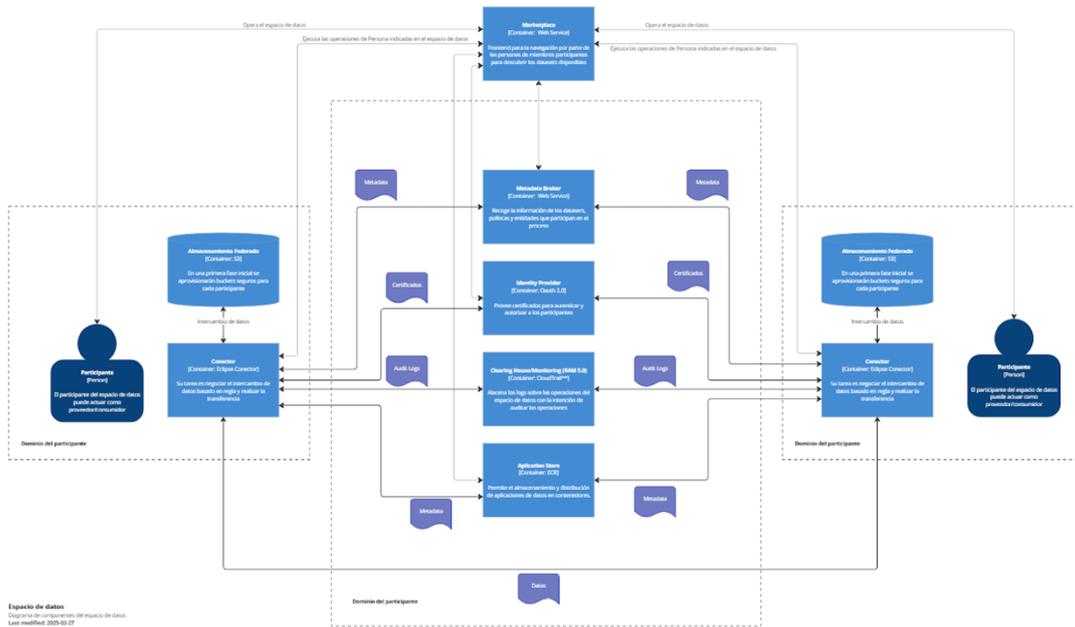
Se han estudiado y explorado otros espacios de datos o alternativas similares de intercambio de la información, aunque los más relevantes son los mencionados anteriormente.

4. Arquitectura del Espacio de Datos

La **arquitectura** descrita en este documento se basa en los principios de descentralización y control soberano de los datos, en plena alineación con la arquitectura de referencia IDS RAM 4.0. Los datos no se centralizan en un único repositorio; por el contrario, cada participante mantiene su propio almacenamiento, asegurando que solo se compartan los datos deseados. El flujo de información entre participantes se realiza a través de un canal seguro, proporcionado por el Data Space Connector (DSC), y todos los datos se encuentran cifrados tanto en tránsito como en reposo.

El **Data Marketplace** proporciona una interfaz web intuitiva que permite a los participantes gestionar tanto sus datos como las políticas de acceso, facilitando así la operativa dentro del espacio de datos. Por su parte, el **Identity Provider (IdP)** garantiza que solo los participantes autenticados puedan acceder a los recursos compartidos, mientras que el **Monitoring Service** asegura que todas las operaciones queden registradas y puedan ser auditadas de forma trazable.

El esquema de arquitectura sería el siguiente:



Veamos a continuación cada uno de los componentes en detalle:

1. **Participantes:** Cada participante en el espacio de datos gestiona su propio almacenamiento y datos. Los datos no se centralizan, sino que permanecen dentro del control de los proveedores de datos; el dominio del participante contiene la capa de almacenamiento que, como se ha comentado, se proveerá en una primera fase a todos los participantes a través de un espacio seguro en S3 para reducir las barreras de entrada.
2. **Personas:** Dentro del espacio de datos tenemos usuarios que necesitan tener visibilidad de la información que se comparte en el espacio de datos para operar los conectores. De hecho, se plantea la creación del Marketplace como un servicio web con frontend para que el usuario pueda operar.
3. **Almacenamiento:** Eventualmente, cada participante contará con su propio almacenamiento desde el que podrá compartir datos de manera soberana. En una primera fase se plantea proveer a los participantes de un espacio seguro en Amazon S3.
4. **Data Space Connector (DSC):** Facilita la integración de diferentes proveedores de datos, gestionando el intercambio de datos entre los diferentes nodos. Actúa como una capa de intermediación de confianza.
5. **Data Marketplace:** Servicio con interfaz web para facilitarle a las personas la operativa dentro del espacio de datos. Dicho marketplace será el responsable de manejar las operaciones del espacio de datos, aunque eventualmente los participantes podrían contar con sus propios frontales web.
6. **Metadata Browser:** Servicio que almacenará y distribuirá la información sobre qué datasets se publican, qué políticas se aplican, etc. y las facilitará al Data Marketplace para su visualización.

7. **Identity Provider:** Emite los certificados y autentica los participantes (tanto las personas a través del Marketplace, cómo los conectores).
8. **Clearing house:** se denominará Monitoring Service para la RAM 5.0 (Reference Architecture Model 5.0) que será la nueva arquitectura propuesta por IDSA para el espacio de datos. Su función es recoger los logs de todas las operaciones que ocurren en el espacio de datos. En una primera fase se plantea el uso de Amazon CloudTrail.
9. **Consumer Data Applications:** Los consumidores de los datos pueden acceder a las aplicaciones de datos a través de APIs seguras y utilizar los contenedores para realizar las transformaciones adecuadas sobre los datos.

4.A. Selección de Proveedores Tecnológicos.

La construcción de un espacio de datos alineado con los principios de interoperabilidad, soberanía y apertura requiere una elección cuidadosa de los componentes tecnológicos y de los proveedores que los soportan. Esta selección debe equilibrar la **adhesión a estándares abiertos**, la **madurez tecnológica**, la **eficiencia operativa** y la **capacidad de evolución futura**. En este contexto, se ha optado por una arquitectura basada en componentes open source impulsados por la **Eclipse Foundation**, integrados sobre servicios gestionados de **Amazon Web Services (AWS)**, con el objetivo de garantizar tanto la agilidad en el despliegue como la compatibilidad con las iniciativas europeas como **Gaia-X**, **IDSA** y **FIWARE**.

Justificación de componentes Eclipse Dataspace y uso de AWS

En el diseño y despliegue de un espacio de datos conforme al modelo de referencia de IDSA, se ha optado por utilizar **Eclipse Dataspace Components (EDC)** como base tecnológica. Esta decisión responde a su creciente consolidación como el conjunto de referencia para habilitar conectores de datos abiertos, interoperables y seguros dentro del ecosistema europeo. EDC está directamente alineado con la arquitectura IDS-RAM y es respaldado por la **Eclipse Foundation**, lo que garantiza su evolución sostenible dentro de una comunidad activa y abierta.

¿Por qué AWS y no otras plataformas como Gaia-X?

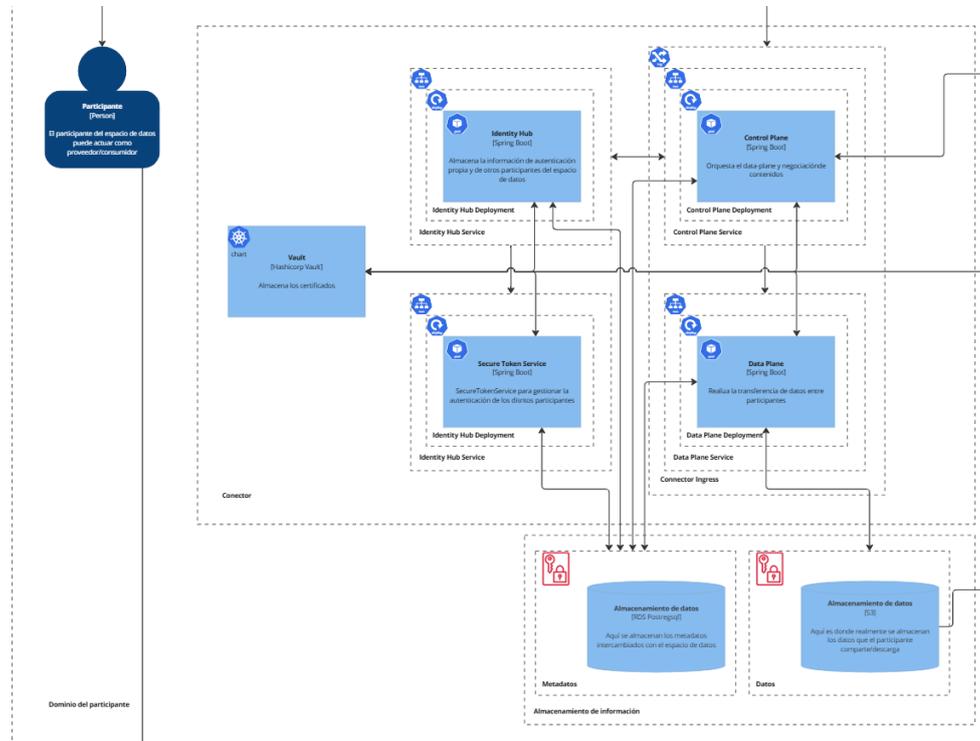
Aunque Gaia-X establece los principios y normas para los espacios de datos europeos (como soberanía, portabilidad y transparencia), no ofrece una infraestructura concreta para su implementación. En cambio, **AWS proporciona los servicios de infraestructura necesarios con alta disponibilidad, escalabilidad y eficiencia operativa**, permitiendo la rápida puesta en marcha de pilotos y entornos productivos. Además, AWS forma parte activa del ecosistema europeo de datos, colaborando con **Eclipse Foundation**, **Gaia-X** y **FIWARE**, lo que refuerza su compatibilidad con iniciativas de soberanía digital.

A continuación se describen los beneficios y el alineamiento de la arquitectura EDME propuesta con el modelo open source:

- 1. Alineamiento con soluciones open source**
El despliegue se apoya en **EKS (Elastic Kubernetes Service)** para orquestación de contenedores, permitiendo la gestión de microservicios open source de forma escalable y automatizada. EKS está basado en Kubernetes, una tecnología completamente abierta y ampliamente adoptada.
- 2. Bases de datos abiertas sobre infraestructura propietaria**
Se utilizan motores como **PostgreSQL**, una base de datos relacional de código abierto, desplegada sobre servicios gestionados de AWS. Esto garantiza compatibilidad con estándares abiertos y evita el bloqueo tecnológico, manteniendo la portabilidad hacia otras nubes si fuese necesario.
- 3. Almacenamiento de objetos (blob)**
Aunque **Amazon S3** es un servicio específico de AWS, es un estándar de facto en la industria, con APIs ampliamente compatibles con soluciones equivalentes en otros proveedores (como Azure Blob Storage o Google Cloud Storage). Esto permite abstraer la lógica de almacenamiento y preservar la interoperabilidad.
- 4. Monitorización y observabilidad**
Dado que las especificaciones de observabilidad en los espacios de datos aún están en desarrollo dentro del entorno IDSA, se han seleccionado herramientas de **bajo coste y probada eficacia dentro de nuestra organización**, priorizando la experiencia interna y la sostenibilidad del despliegue inicial. Se contempla, a futuro, **la migración hacia soluciones open source específicas**, una vez que los estándares de observabilidad estén más maduros y definidos por la comunidad.
- 5. Evolución de AWS hacia componentes abiertos**
AWS ha intensificado su inversión y soporte a proyectos open source, con contribuciones activas a fundaciones como CNCF, Apache y Eclipse. Esta evolución facilita el uso de AWS como plataforma base sin comprometer los principios de apertura y portabilidad del espacio de datos.

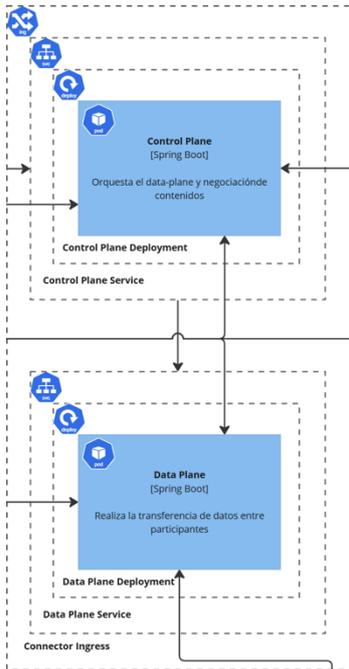
La elección de AWS como plataforma tecnológica no implica un alejamiento de los principios de **soberanía, interoperabilidad y apertura** promovidos por Gaia-X y la IDSA. Por el contrario, la estrategia se fundamenta en un **modelo híbrido**, que aprovecha la madurez tecnológica de AWS, combinado con componentes y estándares open source como EDC, PostgreSQL y Kubernetes, asegurando la flexibilidad y portabilidad del ecosistema. Además, la participación activa de AWS en iniciativas como **Eclipse Dataspaces, Gaia-X y FIWARE** refuerza su compromiso con la construcción colaborativa de una infraestructura europea de datos confiable.

Dominio del Participante



En el dominio del participante podemos encontrar principalmente 3 grandes grupos de componentes:

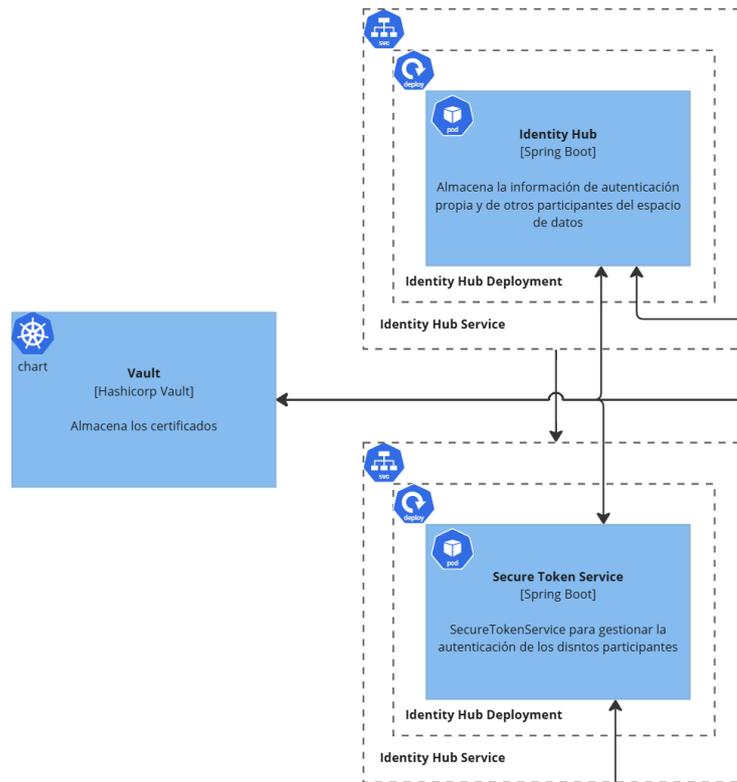
1. **Conector del espacio de datos:** estrictamente hablando el conector está formado por 2 componentes, como se ve en esta imagen:



- Control Plane:** encargado de la negociación de contenido y de la identificación de los participantes, utilizando para ello otros servicios desplegados dentro del dominio del propio participante.
- Data Plane:** es el componente responsable del intercambio de datos entre participantes.

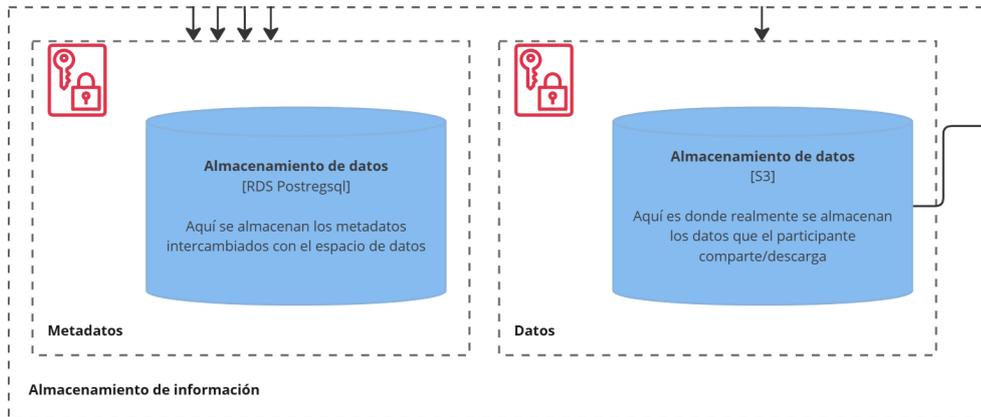
Como puede apreciarse en la imagen, también se detallan el **runtime** a utilizar y los **mecanismos de despliegue**. En este caso, se trata de servicios desarrollados en **Spring Boot y Java**, que serán desplegados en un **clúster de Kubernetes**, incluyendo el **Ingress correspondiente** para permitir que otros participantes puedan comunicarse con dichos servicios de forma segura y eficiente.

2. **Servicios de gestión de la identidad:** se trata de un conjunto de servicios encargados de gestionar la identidad propia, incluyendo el almacenamiento de identificadores propios y de otros participantes, así como de los certificados necesarios para su autenticación y autorización.



En este contexto, se distinguen **tres servicios principales**:

- Identity Hub:** este servicio, utilizado internamente dentro del dominio del participante, es responsable de almacenar los identificadores (IDs) de los participantes del espacio de datos, utilizando para ello DID:WEB generados por el dataspace issuer. También gestiona los identificadores propios del participante. Tal como se muestra en la imagen, se trata de un servicio desarrollado con Spring Boot en Java y desplegado en Kubernetes.
 - Secure Token Service:** este servicio se encarga de la gestión de los web tokens utilizados en la autenticación de los participantes. Al igual que el Identity Hub, es un servicio desarrollado con Spring Boot en Java y desplegado en Kubernetes, como se aprecia en la imagen.
 - HashiCorp Vault:** utilizado en el dominio del participante para el almacenamiento seguro de certificados públicos de otros participantes, así como de los certificados propios. En este caso, se empleará el chart Helm oficial de HashiCorp para su despliegue dentro del clúster de Kubernetes.
3. **Almacenamiento de la información:** finalmente, el participante necesitará almacenar no solo los datos que serán intercambiados en el espacio de datos con otros participantes, sino también los metadatos necesarios para su gestión e intercambio. Esto incluye identificadores, tokens, políticas aplicadas a los datasets, metadatos asociados a los propios conjuntos de datos, entre otros.

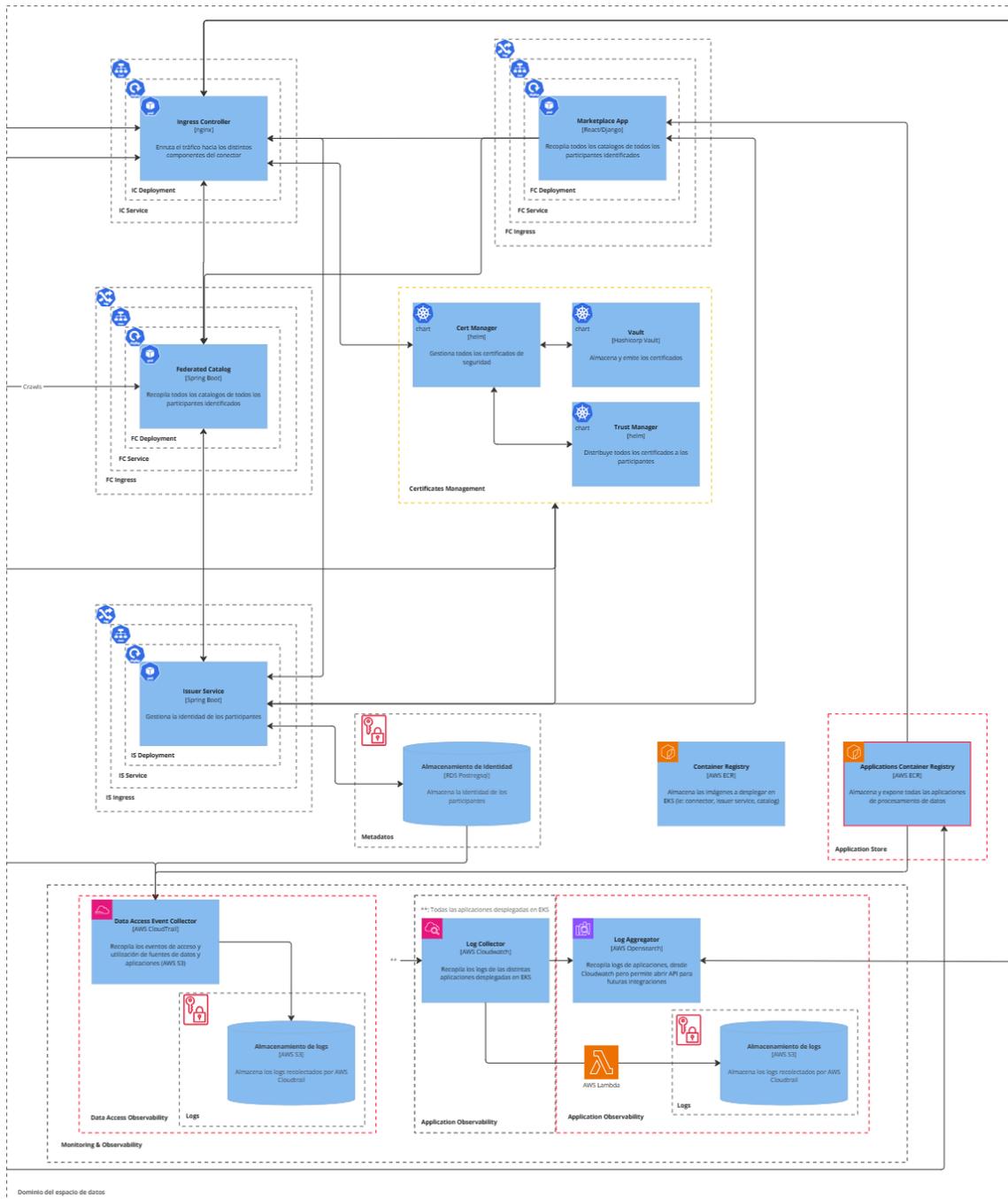


Por ello, se distinguen **dos sistemas de almacenamiento diferenciados**:

1. **Almacenamiento de metadatos:** se empleará una base de datos PostgreSQL, desplegada en Amazon RDS, para el almacenamiento de información relacionada con metadatos: tokens, políticas aplicadas a datasets, metadatos de los propios datasets, entre otros. Esta base de datos estará encriptada en reposo (“at rest”) utilizando claves gestionadas a través de AWS KMS (Key Management Service).
2. **Almacenamiento de datos:** en una primera iteración, se utilizará Amazon S3, lo cual simplificará significativamente la implementación, ya que no será necesario adaptar el data plane a diferentes servicios de almacenamiento. Al igual que en el caso anterior, se aplicará cifrado en reposo mediante KMS.

Con esto, quedan revisados todos los componentes del dominio del participante que serán incluidos en la primera iteración de la arquitectura. A continuación, se detallan los elementos correspondientes al dominio del espacio de datos.

Dominio del Espacio de Datos



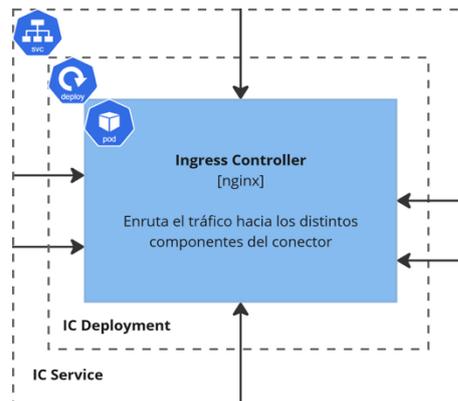
Veremos 7 grandes bloques de servicios que son los que listamos a continuación:

- Ingress Controller
- Marketplace
- Data Catalog
- Identity Services
- Monitoring Services
- Application Store

- Container Registry

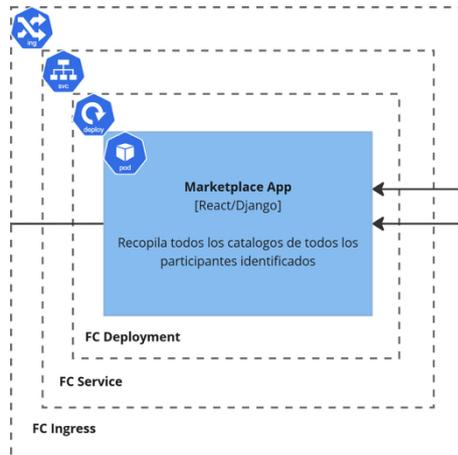
A continuación, detallaremos cada uno de estos bloques.

1. **Ingress Controller:** en este caso se ha seleccionado nginx como ingress controller, lo que nos permitirá acceder a los distintos servicios desplegados en el cluster de kubernetes.

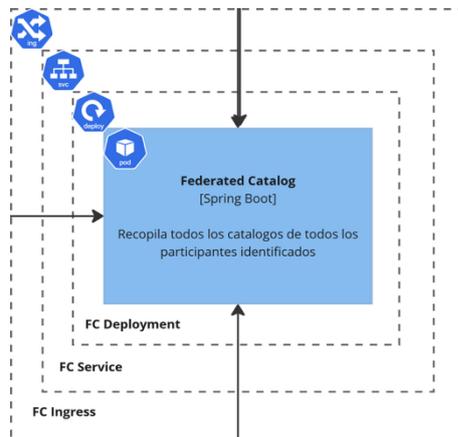


2. **Marketplace:** estrictamente hablando, este componente **no está recogido explícitamente** en la arquitectura de referencia de **IDSA**. No obstante, se considera esencial incluir este **servicio web**, ya que proporcionará la capacidad necesaria para que las personas de cada participante **gestionen datasets, políticas de acceso, etc.** De otro modo, resultaría prácticamente inviable operar en el espacio de datos de forma completamente automática.

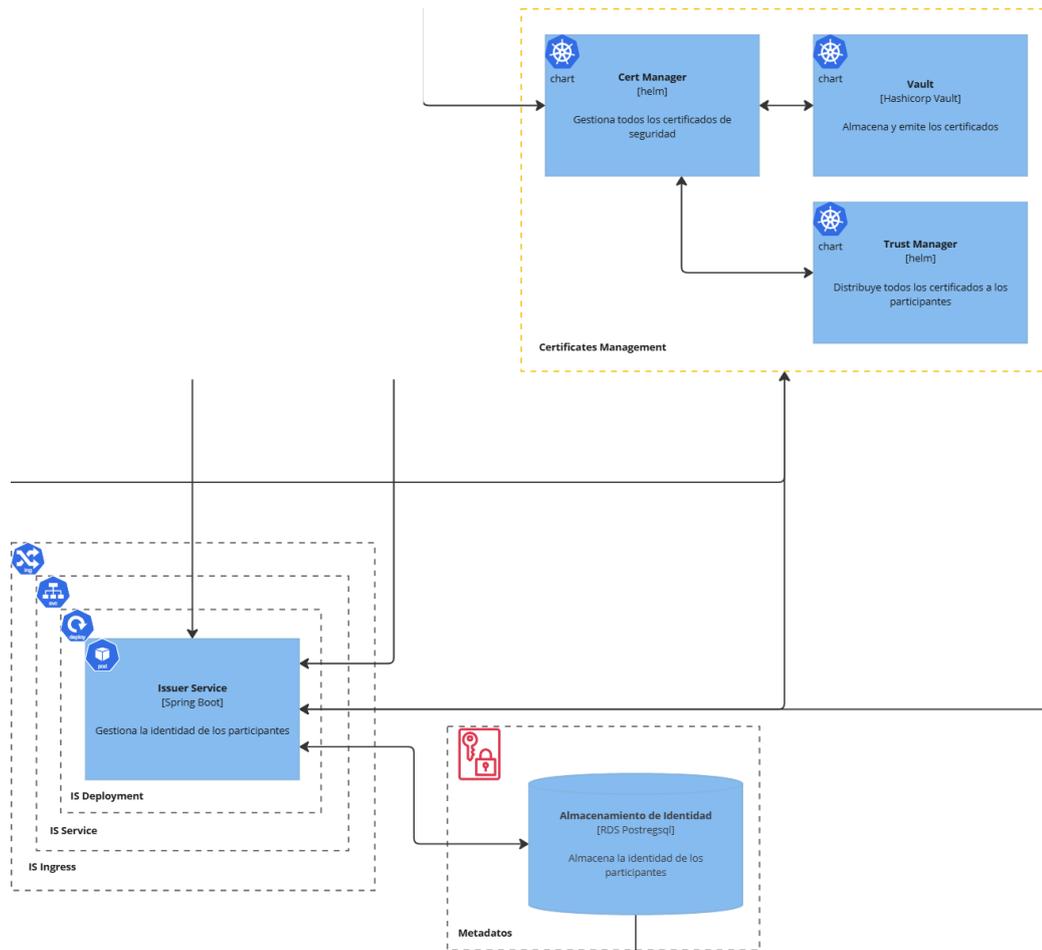
Este **marketplace** estará compuesto por un **front-end** y un **back-end**, así como por distintos **módulos funcionales** que permitirán interactuar con las diversas entidades del espacio de datos. En este caso, y dado que se trata de tecnologías ampliamente conocidas y utilizadas en la organización, se optará por **React.js** para el front-end y **Django (Python)** para el back-end. El servicio será desplegado en **Kubernetes**, incluyendo su correspondiente **ingress** para la gestión del tráfico de red.



3. **Catálogo de Datos:** para facilitar la operación del marketplace, se desplegará un servicio de catálogo federado que permite realizar crawling sobre los distintos participantes y unificar sus catálogos de datos, lo que simplifica y centraliza el acceso a los mismos desde el propio marketplace. Al igual que otros servicios de la arquitectura, este componente será desplegado en Kubernetes, incluyendo su correspondiente Ingress para permitir el acceso externo de forma controlada.



4. **Servicios de Identidad:** En este apartado distinguiremos 2 grandes bloques, gestión de los identificadores de los participantes y gestión de la confianza (autorización y gestión de certificados)

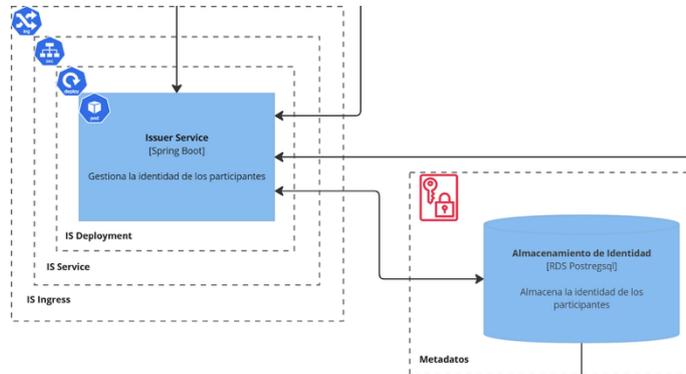


Dentro de estos dos grandes grupos, cabe destacar que la gestión de la confianza podría llevarse a cabo mediante los servicios que Gaia-X proporciona en forma de componentes adaptables y desplegados, diseñados para gestionar las cadenas de certificados y la confianza entre participantes. Sin embargo, en esta primera iteración se ha optado por utilizar **cert-manager**, una solución originalmente desarrollada por RedHat y posteriormente liberada como **proyecto open source** para la comunidad. Actualmente, cert-manager se ha convertido en el **estándar de facto** en la industria para la **gestión de certificados** en entornos **Kubernetes**. Esta elección se complementa con **trust-manager** —proyecto asociado al anterior—, así como con el uso de **HashiCorp Vault** para el almacenamiento seguro de certificados.

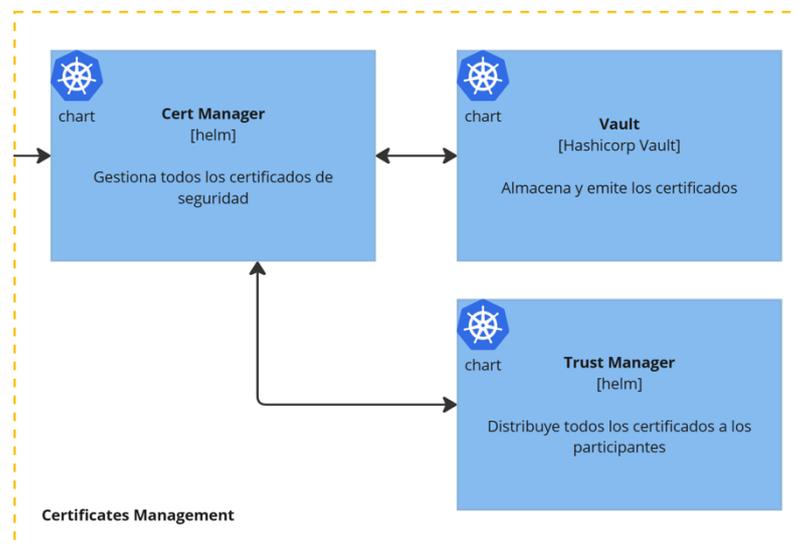
A continuación, se describen **cada uno de los componentes** por separado:

- **Identity Services:** este servicio se encarga de la emisión de identificadores para los distintos participantes. Podría ser gestionado desde el propio marketplace, permitiendo el alta de nuevos participantes a través de una sección de administración. Expone una serie de endpoints que asignan identificadores DID:WEB a cada participante y gestionan

su ciclo de vida. Los identificadores generados se almacenan en una base de datos PostgreSQL, desplegada en Amazon RDS y encriptada en reposo mediante una clave gestionada por KMS.



- **cert-manager:** este servicio es responsable de la emisión y revocación de certificados utilizados por los distintos participantes, además de habilitar el cifrado de las comunicaciones en tránsito mediante TLS. Para la generación y almacenamiento seguro de los certificados, se integrará con HashiCorp Vault.
- **trust-manager:** complemento de cert-manager, este servicio distribuye y gestiona los certificados entre los participantes, facilitando la administración de conjuntos de certificados confiables y automatizando su sincronización en el espacio de datos.

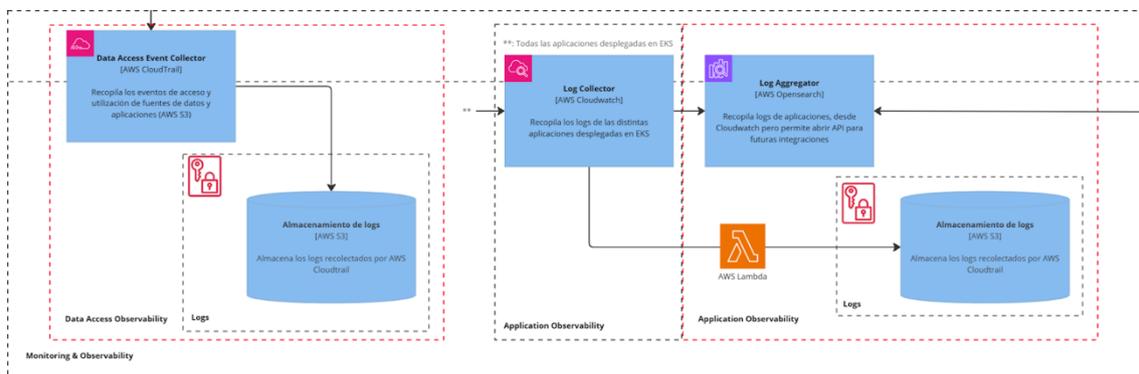


5. **Servicios de monitorización:** aunque la explotación de datos de monitorización por parte de los participantes podría resultar interesante, en esta primera iteración —y dado que se adopta el escenario 1 de integración— no se habilitarán integraciones externas para monitorizar conectores on-premise de los participantes. Todos los componentes del

participante estarán gestionados y desplegados sobre la infraestructura del espacio de datos.

Asimismo, la explotación de logs de monitorización no se contempla para esta primera fase, aunque se valora como una posibilidad viable para iteraciones futuras, en las que se podría proporcionar acceso controlado a dichos registros.

De hecho, se plantean tres subdominios diferenciados de monitorización, que podrían ser implementados progresivamente en futuras iteraciones. En concreto, los subdominios propuestos son:



Veamos cada subdominio por separado:

- **Application Observability (1):** En este subdominio se utilizará AWS CloudWatch para recolectar los logs de las distintas aplicaciones desplegadas en EKS, tanto en los dominios de los participantes como en el del espacio de datos. Esta solución proporciona información suficiente para monitorizar y mantener los servicios, así como para detectar y actuar ante incidencias operativas.
- **Application Observability (2):** No obstante, la solución basada únicamente en CloudWatch dificulta la integración con participantes que utilicen conectores on-premise (escenario de integración 3, no contemplado en la primera iteración). Por ello, se propone como evolución futura la ampliación de la observabilidad con AWS OpenSearch, que permite una integración sencilla con CloudWatch, posibilita la explotación avanzada de datos de monitorización y ofrece interfaces compatibles con entornos on-premise.

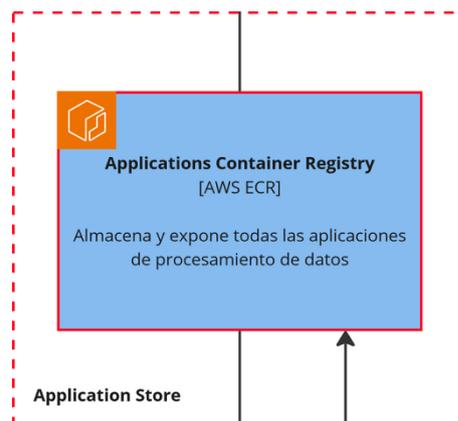
Además, se propone la descarga periódica de logs desde CloudWatch hacia Amazon S3, utilizando AWS Lambda, con el objetivo de:

- **Persistir logs a largo plazo** en almacenamiento cifrado at-rest con KMS.
- **Reducir el volumen de datos almacenados en CloudWatch y OpenSearch**, optimizando así costes y rendimiento.

- **Data Access Observability:** En algunos escenarios, explotar los logs directamente desde la aplicación puede resultar menos eficiente, especialmente al intentar comprender el comportamiento de acceso a datos por parte de los participantes gestionados por el espacio de datos (no aplicable al escenario 3 con participantes on-premise). Para ello, se propone el uso de AWS CloudTrail, que permite auditar el acceso a datos en S3/RDS. Los logs generados también se volcarán en S3 cifrado con KMS, en un bucket separado, para facilitar su almacenamiento y posterior análisis.
6. **Application Store:** Inicialmente, el espacio de datos no contempla el uso de portales ni aplicaciones orientadas al consumo directo de datos, aunque sí admite la publicación y descarga de aplicaciones por parte de los participantes. Estas aplicaciones podrán procesar, transformar o analizar datos, y cada participante podrá cargarlas o descargarlas según su necesidad.

Una de las principales dificultades en este ámbito es que IDSA aún no define un estándar claro sobre cómo deben empaquetarse y distribuirse dichas aplicaciones para que otros participantes puedan utilizarlas de forma interoperable. Por ello, se propone en esta arquitectura el uso de contenedores OCI (como Docker), que permiten empaquetar cada aplicación junto con todas sus dependencias.

No obstante, para que estas aplicaciones puedan utilizarse correctamente no basta con publicar el contenedor: también es necesario documentar su uso, configuración y parámetros de ejecución. En consecuencia, se propone utilizar AWS Elastic Container Registry (ECR) para publicar estos contenedores, y exponerlos junto con su documentación técnica a través del marketplace.



7. **Container Registry:** Como hemos mencionado, muchos servicios del espacio de datos requieren imágenes que sean desplegadas en Kubernetes (en particular en AWS Elastic Kubernetes Service) por ello necesitaremos un repositorio para publicar dichas imágenes, es por ello que usaremos ECR (AWS Elastic Container Registry) de forma similar al

Application Store pero utilizando una instancia separada que nos permite asegurar quien publica y descarga las imágenes de dicho repositorio.



4.C. Integración y Despliegue de la Arquitectura

Antes de adentrarnos en mayor detalle, conviene mencionar los modelos de despliegue de arquitectura e integración entre participantes para comprender mejor los detalles de la arquitectura propuesta.

Se plantean 2 grandes áreas de integración entre participantes:

1. Uso del espacio de datos.
2. Casos de usos construidos sobre el espacio de datos.

Veamos cada uno de esos escenarios por separado.

Integración con el Espacio de Datos

En este apartado nos referimos a cómo un participante entra en el espacio de datos, así como los potenciales escenarios para ejecutar el conector y la infraestructura relacionada necesaria para el correcto funcionamiento junto al resto de participantes. En este aspecto podemos distinguir 3 escenarios:



Espacio de datos
Diagrama de componentes del espacio de datos
Last modified: 2023-02-27

- **[ESCENARIO 1] Infraestructura totalmente gestionada por el espacio de datos:** Este es el escenario elegido para la **primera iteración**. En este caso, el participante no dispone de ningún tipo de infraestructura propia ni conectores on-premise, por lo que toda la infraestructura será proporcionada y gestionada por el promotor del espacio de datos (en este caso, AIE).

Este escenario es el que se abordó en detalle en la **sección sobre arquitectura**.

- **[ESCENARIO 2] Infraestructura parcialmente gestionada por el espacio de datos:** Este escenario deberá resolverse en función de qué componentes tecnológicos posee el participante y cuáles deberían ser proporcionados por el espacio de datos.

Podemos anticipar un caso común: el participante dispone de su propio almacenamiento de datos pero no cuenta con un conector para integrarse en el espacio. En tal caso, sería necesario construir y desplegar un conector cuyo data plane pueda comunicarse con el almacenamiento del participante. Además, se debería poner a disposición una versión de este conector reutilizable por todos los participantes gestionados, tal como se define en el escenario 1.

Este escenario se tratará con mayor profundidad en la sección dedicada a aspectos no funcionales de despliegue e integración.

- **[ESCENARIO 3] Infraestructura no gestionada por el espacio de datos:** En este caso el participante proporciona su propio conector y su propio almacenamiento de datos. El espacio de datos estará preparado para ello, exponiendo los endpoints necesarios que permitan a estos participantes integrarse con los distintos componentes comunes y obtener visibilidad del resto del ecosistema.

Sin embargo, este enfoque requiere que el participante exponga los servicios de su conector, y podría implicar que el espacio de datos deba adaptar ciertos componentes para permitir una integración completa y correcta. Por este motivo, este escenario no se cubre en profundidad en la presente versión de la arquitectura.

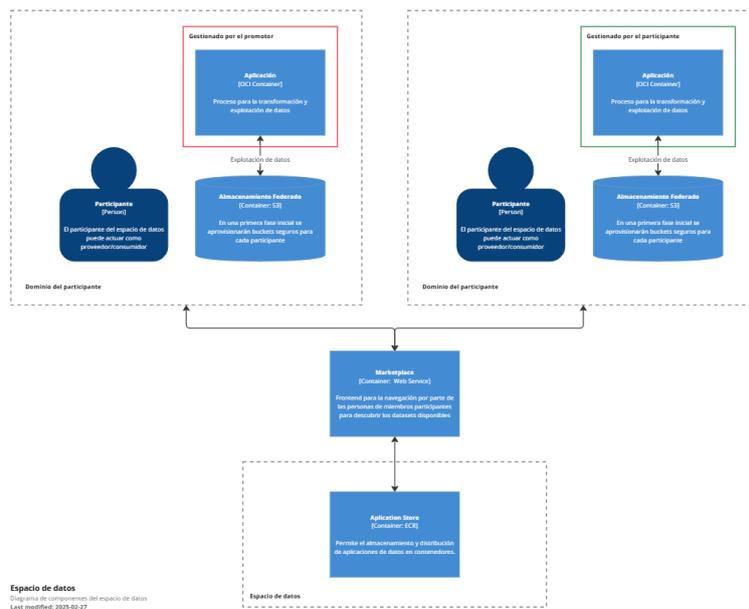
En este contexto, la interacción con el espacio de datos por parte de los usuarios del participante podría ser totalmente opcional, ya que el propio participante podría proveer mecanismos internos que permitan a sus usuarios interactuar con el espacio a través del conector.

Integración de los Casos de Uso

En este contexto, entendemos por caso de uso el conjunto de componentes y procesos que utilizan el espacio de datos para operar sobre los datos y extraer un mayor valor añadido de los mismos.

Es importante destacar que estos procesos y componentes, al cubrir funcionalidades más amplias que las estrictamente definidas por el espacio de datos según la arquitectura de referencia de IDS, no formarían parte oficialmente del espacio de datos, excepto en algunos aspectos sí contemplados por dicha arquitectura, como la publicación y el consumo de aplicaciones de datos (aunque no su ejecución).

Bajo este enfoque, toda manipulación o explotación de datos deberá estar empaquetada como una aplicación, que podrá ser distribuida a través del Application Store, accesible mediante el Marketplace del espacio de datos. En este marco, distinguimos dos escenarios:



- **Infraestructura totalmente gestionada por el promotor del espacio de datos:** Este escenario implicaría el despliegue de procesos que manipulan los datos disponibles, permitiendo que un participante pueda explotarlos sin requerir infraestructura propia.

Por el momento, no se han definido qué tipos de procesos ni qué mecanismos de explotación podrían ser necesarios por parte de los participantes, por lo que este enfoque no se contempla en la primera iteración, excepto para AIE como promotor del espacio de datos.

- **Infraestructura no gestionada por el promotor del espacio de datos:** En este caso, se contempla que el participante se integre con los procesos de publicación y consumo de aplicaciones, sin involucrar al promotor en la ejecución de dichas aplicaciones.

Este escenario sí se abordará para todos los participantes en el contexto de la arquitectura actual.

5. Procesos

En esta sección se detallan los distintos **procesos técnicos necesarios para operar y mantener el espacio de datos**. Se excluyen expresamente los procesos de negocio, como los acuerdos entre participantes o con el promotor del espacio de datos, ya que estos se gestionan en un marco distinto al técnico.

Los procesos se dividen en las siguientes categorías:

1. **Gestión de participantes**
2. **Despliegue de componentes y operaciones en el espacio de datos**
3. **Gestión de aplicaciones y casos de uso**

A continuación, se analiza cada una de estas secciones en detalle.

5.A. Gestión de Participantes

Una vez completados los procesos de negocio necesarios para que un nuevo participante se incorpore al espacio de datos (negociación con el promotor, acuerdos con otros participantes, etc.), comienza la fase de **provisión de acceso técnico al espacio de datos**.

La provisión de infraestructura y su configuración se llevará a cabo siguiendo los principios de **Infraestructure as Code (IaC)**, utilizando herramientas como **Helm** y **Terraform** para crear todos los recursos necesarios que permitan al participante operar dentro del espacio de datos, tal como se detalló en la sección de arquitectura.

Además, se deberá proporcionar acceso al **marketplace** a las personas del nuevo participante. En una primera iteración, se empleará **autenticación básica (usuario y contraseña)**, validada a través de un **correo electrónico corporativo**, verificando que el dominio esté asociado al participante.

En futuras iteraciones, se plantea **reemplazar este sistema por un flujo de autenticación OAuth 2.0**, utilizando un proveedor SSO. En línea con el enfoque open source promovido por IDSA, se consideran soluciones como **Keycloak.org** como posibles proveedores.

5.B. Despliegue de Componentes y Operaciones en el Espacio de Datos

En este apartado se distingue entre la **gestión de la infraestructura** y la **gestión de la configuración**. Ambos aspectos serán gestionados a través de **Git**, adoptando los principios de **GitOps**. Para ello, se utilizarán los **pipelines de Azure DevOps**, permitiendo la automatización de despliegues y el control de cambios.

Esto implica contar, como mínimo, con:

- Un **repositorio para la infraestructura**
- Un **repositorio separado para los componentes y su configuración**

Por ejemplo, la incorporación de un nuevo participante implicaría:

- Modificar la configuración en el repositorio correspondiente
- Desplegar su infraestructura asociada mediante los mecanismos descritos

5.C. Gestión de Aplicaciones y Casos de Uso

Tal como se ha mencionado anteriormente, los casos de uso no forman parte estricta del espacio de datos según la definición de IDSA, aunque sí lo hace el **Application Store**, desde el cual se pueden **publicar, descargar y gestionar aplicaciones** que den soporte a estos casos.

Cuando hablamos de "dar servicio a un caso de uso", nos referimos al despliegue de dichas aplicaciones sobre infraestructura gestionada por el espacio de datos. No obstante, debido a la complejidad y el coste asociado a desplegar aplicaciones individualizadas para cada participante, esta opción queda descartada en una primera iteración.

A futuro, podrían plantearse soluciones que permitan abordar este desafío, aunque requerirían también definir procesos de negocio adicionales (por ejemplo, establecer un modelo de pricing para estas aplicaciones gestionadas).

Por ahora, sin embargo, se contempla la posibilidad de desplegar **casos de uso específicos** (como GDM, Internacional, etc.) **sobre infraestructura gestionada por AIE, utilizando el Application Store** como mecanismo de publicación y distribución de estas aplicaciones.

6. Aspectos no Funcionales

6.A. Seguridad y Privacidad

El espacio de datos está diseñado para **garantizar la soberanía y seguridad de los datos**. Para lograrlo, la **gestión de certificados e identidades de los participantes** es un pilar fundamental.

En esta primera iteración, se ha optado por utilizar los componentes del espacio de datos de Eclipse, ya que están preparados para la comunicación de metadatos en procesos de negociación de contenidos utilizando estándares abiertos. Algunos ejemplos incluyen:

- **ODRL**, para la descripción de políticas de uso de datos.
- **DCAT**, para la expresión de catálogos y datasets.
- **DID:WEB**, para la identificación de participantes.

Además, se han incorporado componentes **estándares industriales consolidados** para la gestión de certificados: **cert-manager**, **trust-manager** y **HashiCorp Vault**

La infraestructura utilizará un **certificado raíz autofirmado (CA)**, que servirá para firmar certificados intermedios. Estos **certificados hijos** serán los que se utilicen para firmar los certificados de cada participante. Esta estrategia permite **revocar certificados individuales comprometidos** sin invalidar el certificado raíz, facilitando la emisión de nuevos certificados sin necesidad de regenerar toda la jerarquía.

Los **certificados** emitidos también se emplearán para **cifrar las comunicaciones mediante TLS (on-transit)**. Para el **cifrado de datos en reposo (at-rest)**, se utilizará **AWS KMS** en los entornos gestionados por el espacio de datos. En el caso de participantes con infraestructura on-premise, se emitirá el certificado necesario para el cifrado en tránsito, pero será responsabilidad del participante demostrar que implementa medidas equivalentes de cifrado at-rest, como garantía para los datos intercambiados con otros participantes.

En cuanto al acceso de usuarios al **Marketplace**, es necesario vincular a cada persona con el participante al que representa. Inicialmente, se emplea **autenticación básica con usuario y contraseña**, complementada por un **proceso de validación por correo electrónico**, que comprobará el dominio del participante.

A **futuro**, se plantea la integración de un **gestor de identidades**, como **Keycloak.org**, que permitiría OAuth 2.0 con autenticación delegada. En este modelo, cada participante proporciona un callback propio, que permitiría configurar el **SSO provider** de forma federada. De este modo,

la **autorización de personas pasaría a estar en el dominio del participante**, eliminando la necesidad de centralizar esta gestión en el espacio de datos.

6.B. Rendimiento y Escalabilidad

En esta fase inicial, y hasta que el valor del espacio de datos se confirme mediante un número significativo de participantes y casos de uso, no se contempla la prestación de servicios en alta disponibilidad ni con escalado dinámico completo.

Sin embargo, se preparará la infraestructura para permitir una evolución futura hacia modelos escalables, implementando:

- **Node groups en Kubernetes** con capacidad de **escalado automático**.
- Separación de node groups según el tipo de servicio y sus necesidades específicas.

En concreto, se distinguen los siguientes grupos dentro de cada espacio de datos:

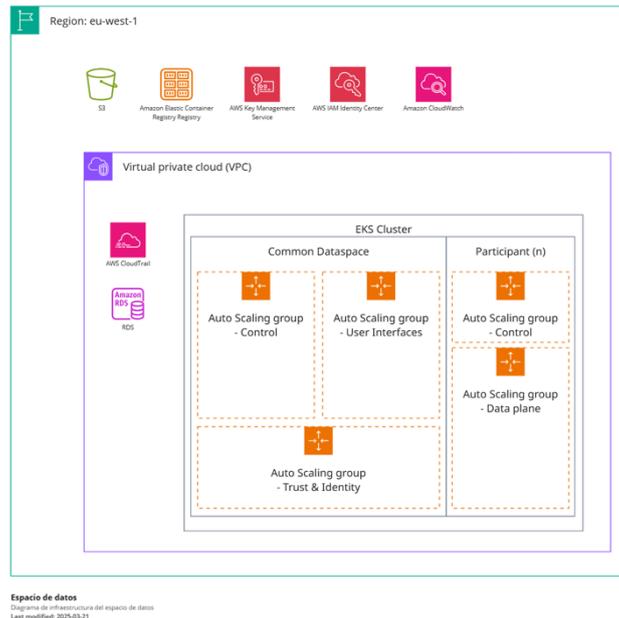
Common Dataspace

- **Control Group:** servicios core del espacio de datos (catálogo, identity issuer, ingress controller).
- **User Interfaces:** interfaz de usuario del Marketplace.
- **Trust & Identity:** servicios relacionados con la gestión de certificados e identidad.

Participant Dataspace

- **Control Group:** servicios del participante, excepto el plano de datos.
- **Data Plane Group:** servicios responsables de las transferencias de datos. Dado su mayor impacto en términos de carga y concurrencia, este grupo requerirá **mayores capacidades de escalabilidad**.

Con esta segmentación se garantiza una base sólida que permitirá adaptar la infraestructura a medida que aumente la carga y el uso del espacio de datos.



6.C. Despliegue e Integración

Para implementar con éxito el espacio de datos, es fundamental **minimizar la intervención humana en su gestión**, garantizando que las operaciones realizadas por los participantes **escalen de forma eficiente y sostenible**, incluso en caso de incidencias con los componentes.

Para ello, se adoptarán los principios de **GitOps** e **Infrastructure as Code (IaC)** tanto para la gestión de la infraestructura como para la configuración de los componentes. Se utilizará **Terraform** para describir la infraestructura y la configuración de forma **declarativa**, mediante ficheros versionables que permitan tratarlos como código. Estos ficheros serán ejecutados mediante **Azure DevOps**, lo que asegurará que la infraestructura y los servicios estén **desplegados correctamente** y con la configuración deseada.

Aunque Kubernetes ya se encarga de mantener la disponibilidad de los pods a través de sus mecanismos de control de despliegue y servicios, sólo mediante la ejecución de los scripts de Terraform se garantiza la consistencia global entre la infraestructura desplegada y su definición declarativa.

Por esta razón, en el futuro se valorará la incorporación de herramientas como **Argo CD**, que permiten verificar y aplicar automáticamente la configuración declarada con una frecuencia determinada. Esta funcionalidad sería especialmente útil para **servicios no gestionados por Kubernetes**, como **RDS, KMS, S3 o CloudWatch**. Para ello, **Argo CD podría desplegarse dentro del propio clúster de Kubernetes**, integrando su operación con el resto de componentes.

6.D. Mantenimiento y Soporte

Como se ha indicado anteriormente, este primer documento establece únicamente los **requisitos mínimos necesarios para una versión funcional del espacio de datos**, cumpliendo con **estándares de seguridad** que aseguran el cumplimiento de los principios fundamentales de operación.

Dentro de estos mínimos, por ejemplo, no se ha adoptado la infraestructura propuesta por Gaia-X (como el *clearing house*) para la gestión de la confianza entre participantes. Sin embargo, en futuras iteraciones se podría considerar reemplazar la solución basada en cert-manager por una integración completa con los servicios de clearing house de Gaia-X.

Asimismo, este documento **no contempla aún la creación de casos de uso específicos**, los cuales se desarrollarán progresivamente una vez que la infraestructura del espacio de datos esté consolidada.

7. Pruebas de Concepto para la validación de la arquitectura propuesta

Cabe destacar que **se han realizado pruebas de concepto**, localmente o incluso en cuenta de pruebas de AWS, **para validar el despliegue** y la interacción entre los componentes propuestos en la arquitectura. Se ha tenido en cuenta no sólo la primera iteración sino también algunos escenarios más avanzados (por ej. cert-manager, keycloak para la introducción de SSO providers en el marketplace, etc)

8. Beneficios de la Arquitectura en IDS

- **Control de Soberanía de los Datos:** Las organizaciones tienen control total sobre sus datos, compartiéndolos sólo bajo condiciones que ellas mismas definen.
- **Cumplimiento Normativo:** Se facilita el cumplimiento de regulaciones de privacidad y protección de datos como el **GDPR**.
- **Seguridad y Confianza:** A través de mecanismos de autenticación, autorización y cifrado, la seguridad de los datos se asegura en todo momento.
- **Escalabilidad y Flexibilidad:** La arquitectura permite añadir nuevos participantes o servicios sin comprometer la privacidad o la soberanía de los datos.

9. Roadmap para la Implementación Práctica de EDME

Hasta el momento hemos descrito la fundamentación técnica para el desarrollo de EDME y a continuación mostraremos el itinerario general que habrá que seguir para su implementación operativa. Un aspecto relevante será la implicación en el proceso de los principales actores de la industria de la música española desde la fase inicial de diseño (enfoque multiactor). Se trata de una iniciativa colaborativa para la compartición de datos y por lo tanto el alcance de acuerdos comunes en aspectos clave será fundamental desde las primeras fases.

Primeramente, se habrán de **debatir conjuntamente diferentes temas fundamentales** que permitan tomar decisiones **para el diseño de EDME**, incluyendo:

- **Estructura de EDME** : centralizada, federada o totalmente descentralizada
- **Autoridad de gobierno de EDME** (generalmente denominada DSGA, *Data Space Governance Authority*): centralizada en una entidad independiente como AIE, o descentralizada en diferentes actores
- **Definición de los requisitos y atributos principales a cumplir por los participantes de EDME**: por ej. que pertenezcan al sector de la música, sean socios de alguna entidad de gestión de derechos musicales, experiencia de más de cinco años en el sector, una determinada capacidad tecnológica o que disponga de algún tipo de certificación
- **Método de verificación de los requisitos**: a través de medios internos de EDME o bien mediante la implicación de entidades terceras como por ejemplo certificadoras
- **Tipos de datos y casos de uso en EDME**: este aspecto es clave para la sostenibilidad de EDME puesto que el impacto de los primeros casos de uso será fundamental para el reclutamiento de nuevos participantes y el futuro desarrollo del espacio
- **Modelo de financiación de EDME**: como se espera hacer sostenible financieramente el espacio EDME, por ejemplo a través de cuotas de socios mediante la venta de conectores etcétera

Una vez que se han tomado estas decisiones de diseño **se planifican los diferentes elementos funcionales de EDME** :

- Reglas : que comportamientos y habilidades se requieren para participar
- Políticas: las reglas de participación expresadas y verificadas en políticas. También se habrán de incluir normativas sectoriales o regulaciones europeas y nacionales relacionadas con datos como Data Governance Act, Data Act o GDPR
- Certificación de pertenencia a EDME: qué mecanismos se utilizan para verificar la participación en EDME

- Registro de participantes: listado de participantes en EDME
- Sistema de identificación: centralizado o descentralizado
- Catálogo de datos: una central, varios federados o catálogos individuales descentralizados

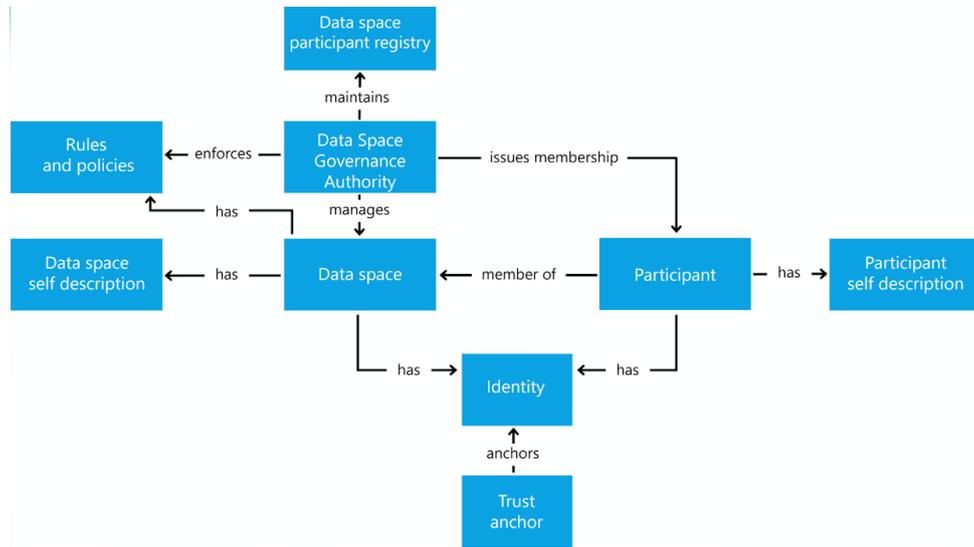


Figura. Entidades en el ED (tomada del IDSA Rulebook, <https://docs.internationaldataspaces.org/ids-knowledgebase/idsa-rulebook>)

Todos estos aspectos se desarrollarán dentro de las **políticas y reglas de EDME** (*membership policies*) que estarán accesibles a cualquier potencial participante para valorar su futura implicación en la iniciativa. Este documento público que recoge las políticas del ED se conoce como *Data Space Self Description* (DSSD). La autoridad de gobierno del ED, DSGA, será la responsable de implementar y hacer seguimiento del cumplimiento reglas y políticas, así como gestionar todo el ciclo de vida de la membresía (validación de requisitos, expedición, control y revocación de credenciales).

Una vez establecidas las políticas y la autoridad de gobierno, **cualquier organización interesada en participar en EDME podrá solicitarlo** a través del procedimiento de registro establecido. Para ello, tendrá que aportar la información y documentación requerida en los formatos indicados, poniendo especial énfasis en las pruebas solicitadas para verificar el cumplimiento de los requisitos mínimos para formar del espacio (se conoce como *Participant self description*). Si todo está correcto, la autoridad DSGA le concederá las credenciales y el participante podrá tener acceso a las diferentes funcionalidades y conjuntos de datos disponibles en EDME.

Una vez dentro de EDME el participante podrá **consultar el catálogo de datos disponible**. Cabe resaltar que determinados proveedores de datos podrán establecer requisitos adicionales a los

consumidores de datos para tener acceso (por ej. que sean de determinados países, que pertenezcan a una determinada industria, o que posean una determinada certificación). Estas son las denominadas **políticas de acceso**. Asimismo, los consumidores de datos también podrán poner sus propias políticas para visibilizar únicamente aquella oferta de datos que los cumpla, como por ej. datos que tengan un determinado sello de su origen.

Cuando un participante muestre interés por un determinado conjunto de datos se abrirá una **negociación entre el proveedor de los datos y el consumidor**. Durante dicha negociación se podrán establecer determinadas **políticas de uso**, como por ejemplo, el tratamiento de datos anonimizados, el uso de determinados sistemas de encriptación, la posibilidad de compartir los resultados con terceras partes de una manera agregada, o definir un determinado periodo de uso.

Con el fin de eficientar las operaciones generales de EDME descritas previamente, se tratará de automatizar al máximo dichos procesos, minimizando aquellos que impliquen la mediación humana.

Asimismo, dentro del ED se puede disponer de un **agente observador** que monitorice la actividad en el espacio y sirva de apoyo para verificar las diferentes reglas y políticas establecidas en los diferentes niveles (*clearing house*).

Finalmente, resaltar que IDSA describe otra serie de **funciones opcionales** que pueden estar presentes en los ED, como por ejemplo el papel de intermediarios o actores que proveen servicios de valor añadido. Entre ellos destacamos el desarrollo de Marketplace, servicios analíticos avanzados o de procesamiento de datos.

10. Conclusión

La arquitectura propuesta para un **espacio de datos internacional**, basado en los principios de la **International Data Spaces Association (IDSA)**, proporciona una solución integral y altamente segura para el intercambio de datos entre organizaciones. Esta solución no solo está alineada con los estándares y mejores prácticas de IDSA, sino que también aprovecha al máximo las capacidades de la nube, utilizando servicios de **Amazon Web Services (AWS)**, lo que garantiza escalabilidad, flexibilidad y alta disponibilidad.

Una de las principales fortalezas de esta arquitectura es su enfoque en la **soberanía de los datos**. En lugar de centralizar los datos en una ubicación única, la arquitectura permite que cada organización mantenga el control total sobre sus datos, asegurando que solo se compartan con aquellas entidades que el propietario de los datos haya autorizado explícitamente. Este enfoque de descentralización es fundamental para garantizar que las

organizaciones no pierdan el control sobre sus activos más valiosos: los datos. Además, las políticas de acceso y las reglas de compartición pueden ser completamente personalizadas y ajustadas según las necesidades y requisitos específicos de cada organización.

El diseño de la **capa de confianza** a través del **Data Space Connector (DSC)** permite la integración y el intercambio de datos entre los participantes de manera **segura y transparente**. La autenticación y autorización estrictas a través de un **Identity Provider (IdP)**, junto con los protocolos de seguridad como **OAuth 2.0**, aseguran que solo las entidades autorizadas puedan acceder a los datos, minimizando riesgos de acceso no autorizado o uso indebido de la información. Además, la **encriptación** de los datos, tanto en tránsito como en reposo, garantiza que los datos estén protegidos en todo momento, cumpliendo con las más estrictas regulaciones de privacidad y protección de datos, como el **GDPR**.

Otro aspecto clave es la **interoperabilidad** que se logra mediante el uso de servicios estándar de AWS como **S3** para almacenamiento y **CloudTrail** para auditoría. Esta interoperabilidad no solo facilita la integración de diferentes fuentes de datos, sino que también asegura que las organizaciones puedan interactuar de manera fluida y sin barreras tecnológicas. Los participantes pueden operar dentro de este espacio de datos de forma sencilla y segura, gracias a la creación del **Data Marketplace**, una plataforma de fácil acceso y uso que facilita la visualización, el acceso y el manejo de los datos, sin la necesidad de contar con conocimientos técnicos avanzados.

La arquitectura propuesta también ha sido diseñada para ser escalable y flexible. Con el tiempo, el espacio de datos puede expandirse sin comprometer la seguridad o la privacidad de los datos, lo que permite agregar nuevos participantes o integrar servicios adicionales según sea necesario. Este diseño modular asegura que el sistema pueda evolucionar de acuerdo con las necesidades cambiantes del mercado y los avances tecnológicos, adaptándose a nuevas demandas y desafíos.

Además, la arquitectura permite a las organizaciones cumplir con las normativas globales de privacidad y seguridad, proporcionando un marco de **gobernanza de datos** claro y auditable. Los mecanismos de monitoreo y auditoría garantizan que todas las actividades dentro del espacio de datos sean transparentes y trazables, lo que aumenta la confianza de los participantes y facilita el cumplimiento de las regulaciones internacionales.

En resumen, esta arquitectura para un espacio de datos internacional ofrece un modelo **seguro, soberano, escalable y flexible** que no solo facilita el intercambio de datos entre organizaciones, sino que también protege los derechos y el control de los propietarios de

los datos. Al alinearse con los principios de **International Data Spaces Association (IDSA)** y aprovechar las capacidades de **AWS**, esta solución proporciona un entorno confiable y de confianza para la colaboración entre múltiples actores, sin comprometer la seguridad, la privacidad o la soberanía de los datos. La propuesta no solo busca satisfacer las necesidades actuales del intercambio de datos, sino también anticipar las necesidades futuras, haciendo que el espacio de datos sea una plataforma adaptable y evolucionable.