ARQUITECTURA DEL ESPACIO DE DATOS

EDS4AGRO

Versión	Fecha	Descripción
1.0	26/03/2026	Versión inicial del documento

Financiado por la Unión Europea
NextGenerationEU
MINISTERIO PARA LA TRANSFORMACIÓN DIGITAL Y DE LA FUNCIÓN PÚBLICA
SECRETARÍA DE ESTADO DE DIGITALIZACIÓN E INTELIGENCIA ARTIFICIAL
Plan de Recuperación, Transformación y Resiliencia

Contenido

Financiado por la Unión Europea NextGenerationEU

MINISTERIO PARA LA TRANSFORMACIÓN DIGITAL Y DE LA FUNCIÓN PÚBLICA

SECRETARÍA DE ESTADO DE DIGITALIZACIÓN E INTELIGENCIA ARTIFICIAL

Plan de Recuperación, Transformación y Resiliencia

INTRODUCCIÓN

El presente documento tiene por objeto describir la arquitectura del espacio de datos EDS4Agro, proporcionando una visión estructurada de los elementos técnicos que soportan su funcionamiento y de las relaciones existentes entre sus componentes. Se concibe, además, como un complemento al marco de gobernanza definido en el Libro de Reglas y en el conjunto de políticas del espacio de datos, ofreciendo una perspectiva técnica que permite entender cómo se implementan, a nivel de arquitectura, los principios de interoperabilidad, seguridad, soberanía del dato y control del uso de la información.

La arquitectura de EDS4Agro se basa en un enfoque federado y distribuido, en el que los participantes mantienen el control sobre sus datos, estableciendo mecanismos de intercambio basados en estándares abiertos y en el uso de conectores que garantizan la gestión de accesos y la aplicación de políticas de uso.

El alcance del documento se limita a la descripción de alto nivel de la arquitectura del sistema, incluyendo sus principales componentes, los flujos de intercambio de datos y los mecanismos de identidad, seguridad y trazabilidad. No se pretende proporcionar un nivel de detalle técnico exhaustivo, sino una visión clara y coherente que facilite la comprensión del funcionamiento del espacio de datos desde un punto de vista arquitectónico.

VISIÓN GENERAL DEL ESPACIO DE DATOS

El espacio de datos EDS4Agro constituye un ecosistema digital orientado a facilitar el intercambio seguro, controlado e interoperable de datos dentro del sector agroalimentario. Su objetivo es habilitar un entorno de colaboración entre entidades públicas y privadas, permitiendo la generación de valor a partir del dato, manteniendo en todo momento la soberanía sobre la información compartida.

El ecosistema está compuesto por distintos tipos de participantes que interactúan entre sí, entre los que se incluyen entidades proveedoras de datos y servicios, entidades consumidoras y el operador del espacio de datos, encargado de la provisión de los servicios comunes y la coordinación técnica del entorno. Cada participante mantiene el control sobre sus propios datos, definiendo las condiciones bajo las cuales estos pueden ser compartidos y utilizados por terceros.

EDS4Agro se basa en un modelo de arquitectura federada y distribuida, en el que los datos no se centralizan, sino que permanecen en los sistemas de los participantes. El espacio de datos actúa como facilitador del intercambio, proporcionando los mecanismos necesarios para el descubrimiento de activos, la gestión de condiciones de acceso y el establecimiento de relaciones de intercambio entre las partes.

El funcionamiento del espacio de datos se articula en torno a un proceso estructurado de intercambio, que comprende la publicación de activos de datos por parte de los proveedores, su descubrimiento a través de mecanismos de catálogo, la negociación de las condiciones de acceso mediante contratos de datos y, finalmente, el acceso o transferencia de la información conforme a dichas condiciones.

ARQUITECTURA DEL ESPACIO DE DATOS

Vista general de la arquitectura

La arquitectura del espacio de datos EDS4Agro se basa en un enfoque distribuido y federado, en el que los distintos participantes mantienen el control sobre sus datos y servicios, mientras que el espacio de datos proporciona los mecanismos necesarios para garantizar el intercambio seguro, interoperable y gobernado de la información.

Financiado por la Unión Europea NextGenerationEU
MINISTERIO PARA LA TRANSFORMACIÓN DIGITAL Y DE LA FUNCIÓN PÚBLICA
SECRETARÍA DE ESTADO DE DIGITALIZACIÓN E INTELIGENCIA ARTIFICIAL
Plan de Recuperación, Transformación y Resiliencia

Desde una perspectiva global, la arquitectura se organiza en varios bloques funcionales que, de manera conjunta, permiten habilitar el ciclo completo de compartición de datos dentro del ecosistema. Estos bloques incluyen los servicios centrales del espacio de datos, los nodos de los participantes y el conjunto de componentes que soportan el intercambio, la interoperabilidad y la gestión del dato.

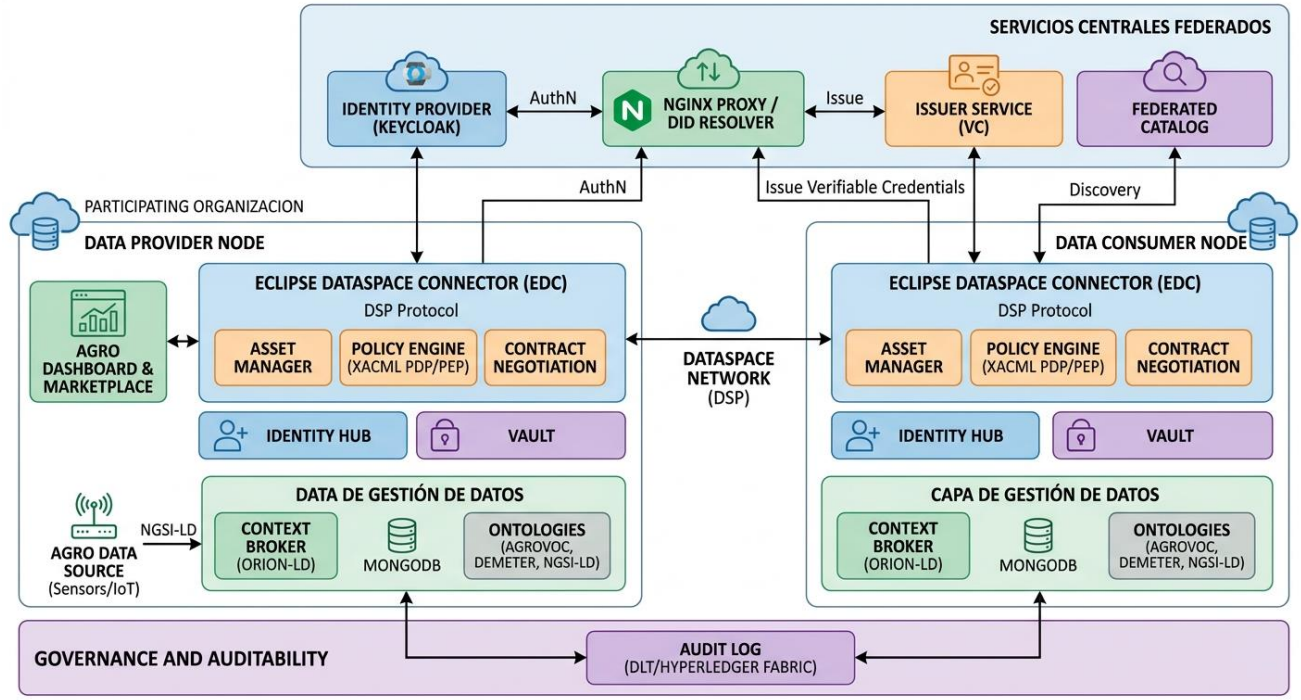
En primer lugar, el espacio de datos dispone de una capa de **servicios centrales**, encargada de proporcionar los mecanismos de confianza, identidad y gobernanza. Estos servicios actúan como elementos habilitadores comunes para todos los participantes, permitiendo la autenticación, la gestión de credenciales, la validación de políticas de acceso y la trazabilidad de las transacciones. Su papel es fundamental para garantizar que únicamente entidades autorizadas puedan participar en el ecosistema y acceder a los datos bajo condiciones previamente definidas.

En segundo lugar, cada participante del espacio de datos opera a través de un **nodo propio**, que constituye la unidad básica de integración en el ecosistema. Estos nodos incluyen los componentes necesarios para exponer, consumir y gestionar datos, manteniendo en todo momento la soberanía sobre la información. A través de estos nodos, los participantes publican sus activos de datos, definen las condiciones de acceso y establecen relaciones de intercambio con otros actores del espacio de datos.

El intercambio de información entre participantes se articula mediante un **plano de datos e interoperabilidad**, que define tanto los mecanismos técnicos de intercambio como los modelos de datos y estándares utilizados. En este contexto, el uso de conectores de espacio de datos permite implementar funcionalidades clave como la publicación de catálogos, el descubrimiento de activos, la negociación de contratos y la transferencia de datos, siguiendo principios de soberanía y control descentralizado.

Adicionalmente, la arquitectura incorpora una capa de **gestión del dato y contexto**, que permite estructurar, almacenar y acceder a la información de manera coherente y en tiempo real. Esta capa facilita la interoperabilidad semántica mediante el uso de estándares como NGSI-LD y JSON-LD, asegurando que los datos compartidos no solo sean accesibles, sino también comprensibles por todos los participantes del ecosistema.

En general, la arquitectura del espacio de datos queda reflejada en la siguiente figura.



Componentes principales

A continuación, se describen los principales componentes y aspectos que conforman el espacio de datos.

Servicios de identidad y confianza

El espacio de datos EDS4Agro incorpora un conjunto de servicios de identidad y confianza que constituyen un elemento fundamental de su arquitectura, permitiendo garantizar la identificación segura de los participantes, la autenticación de sus interacciones y el establecimiento de relaciones de confianza dentro del ecosistema. Estos servicios se diseñan conforme a un enfoque híbrido que combina mecanismos tradicionales de gestión de identidades y accesos (IAM) con modelos de identidad digital descentralizada (Self-Sovereign Identity, SSI), proporcionando tanto control centralizado de acceso a servicios como soberanía del participante sobre su identidad digital.

En primer lugar, la arquitectura integra un **proveedor de identidad**, encargado de gestionar los procesos de autenticación y autorización inicial de los participantes en el acceso a los servicios del espacio de datos. Este componente permite la gestión de identidades, credenciales de acceso, roles y sesiones, así como la emisión de tokens de autenticación (por ejemplo, basados en estándares como OAuth2/OpenID Connect), que son utilizados para el acceso seguro a los distintos componentes del sistema.

Complementariamente, el sistema incorpora un **servicio emisor de credenciales verificables (Issuer)**, responsable de la generación y firma de credenciales digitales asociadas a los participantes. Estas credenciales contienen atributos relevantes (como identidad organizativa, rol dentro del espacio de datos o capacidades técnicas) y se estructuran conforme a modelos de credenciales verificables, permitiendo su verificación por terceros sin necesidad de intermediación directa del emisor.

La arquitectura adopta un modelo de **identidad digital descentralizada (SSI)**, en el que cada participante dispone de uno o varios **identificadores descentralizados (DID)**, que actúan como identificadores únicos y resolubles dentro del ecosistema. Estos identificadores se gestionan de forma autónoma por cada entidad, permitiendo la creación de identidades digitales sin dependencia de autoridades centralizadas y facilitando la interoperabilidad entre distintos dominios.

En este contexto, los participantes pueden presentar **credenciales verificables (VC)** y **presentaciones verificables (VP)** durante los procesos de autenticación y autorización, permitiendo demostrar atributos específicos sin necesidad de revelar información adicional no necesaria (principio de minimización de datos). Este mecanismo resulta especialmente relevante en escenarios de intercambio de datos, donde es necesario verificar condiciones de acceso sin comprometer la privacidad de los participantes.

La combinación de estos elementos permite establecer un **modelo de confianza distribuido**, en el que la identidad de los participantes, sus atributos y sus capacidades pueden ser validados de forma criptográficamente verificable. Este modelo se integra con los mecanismos de control de acceso del espacio de datos, permitiendo que las decisiones de autorización se basen tanto en la identidad como en los atributos certificados de los participantes.

Asimismo, estos servicios se integran de forma transversal con el resto de la arquitectura, en particular con los conectores de espacio de datos y los mecanismos de aplicación de políticas, garantizando que todas las interacciones dentro del ecosistema se realizan sobre la base de identidades verificadas y bajo condiciones de acceso controladas.

Framework de gobernanza y control de acceso

La arquitectura de EDS4Agro incorpora un framework específico de gobernanza y control de acceso cuyo objetivo es asegurar que el acceso a los datos y servicios del ecosistema se produce de forma controlada, verificable y conforme a las condiciones definidas por sus titulares. Este framework constituye una capa esencial de la arquitectura del espacio de datos, ya que traduce los principios de soberanía del dato, control del uso y confianza entre participantes en mecanismos técnicos capaces de aplicarse durante la operación real del sistema.

Desde el punto de vista arquitectónico, este framework no se limita a autenticar participantes, sino que añade una capa adicional de decisión sobre el acceso a los recursos. De este modo, la arquitectura diferencia claramente entre, por un lado, la identificación y autenticación de una entidad participante y, por otro, la determinación de si esa entidad puede o no realizar una acción concreta sobre un recurso determinado. Esta separación es importante porque permite implementar políticas de acceso más granulares y alineadas con el modelo de gobernanza del espacio de datos.

En EDS4Agro, el framework de gobernanza y control de acceso se apoya principalmente en dos pilares complementarios. El primero de ellos es el uso de las capacidades de control de acceso y negociación propias de los conectores del espacio de datos, que permiten vincular activos, políticas y contratos. El segundo es la incorporación de un marco adicional de control basado en XACML, que proporciona una capa de autorización de grano fino sobre recursos y acciones concretas. Esta combinación permite pasar de un modelo de acceso basado únicamente en la existencia de una relación contractual o una credencial válida, a un modelo más completo en el que también pueden evaluarse atributos, roles, acciones y condiciones específicas de uso.

El uso de XACML responde a la necesidad de expresar políticas de autorización de manera formal, estructurada y evaluable automáticamente. XACML permite implementar un modelo de control de acceso basado en atributos (ABAC, Attribute-Based Access Control), en el que las decisiones no dependen exclusivamente de identidades estáticas o listas predefinidas de permisos, sino de la evaluación de atributos del sujeto, del recurso y de la acción solicitada. En el contexto de EDS4Agro, esto resulta especialmente útil porque el acceso a un activo puede depender no solo de quién solicita el acceso, sino también del tipo de participante, de su rol, del recurso concreto al que intenta acceder, del método utilizado o de otras condiciones definidas en las políticas del espacio de datos.

Arquitectónicamente, este framework incorpora los componentes clásicos del ecosistema XACML. En primer lugar, un repositorio o archivo de políticas, que almacena las políticas de control de acceso definidas para los distintos recursos. En segundo lugar, un Punto de Administración de Políticas (PAP), a través del cual pueden definirse, mantenerse y actualizarse dichas políticas. En tercer lugar, un Punto de Decisión de Políticas (PDP), responsable de evaluar las solicitudes de acceso frente a las políticas vigentes y emitir un veredicto final, que puede materializarse como permitir, denegar o considerar no aplicable una determinada solicitud. Esta organización permite desacoplar la definición de las reglas de acceso de su evaluación operativa, facilitando tanto la gobernanza como la auditabilidad del sistema.

Una de las aportaciones más relevantes de este enfoque en EDS4Agro es que el control de acceso puede aplicarse sobre un triplete formado por sujeto, acción y recurso. Esto permite, por ejemplo, diferenciar no solo qué participante puede acceder a un determinado activo, sino también qué operaciones concretas puede realizar sobre él. Desde una perspectiva arquitectónica, este aspecto es especialmente valioso porque evita modelos excesivamente binarios de acceso y permite incorporar controles más ajustados a los requisitos reales del ecosistema. Así, una entidad puede estar autorizada para consultar un recurso mediante una operación determinada, pero no necesariamente para realizar otras acciones sobre el mismo endpoint o activo.

Este framework de gobernanza y control de acceso se integra con el modelo de identidad y confianza descrito anteriormente. En particular, las decisiones de autorización pueden apoyarse tanto en la identidad autenticada del participante como en los atributos acreditados mediante credenciales verificables o en la información contenida en tokens de autenticación emitidos por el proveedor de identidad. De esta forma, el sistema articula una cadena de confianza en la que la identidad, las credenciales, las políticas y la decisión de acceso forman parte de un mismo flujo lógico. Esta integración es fundamental para garantizar que el acceso al dato no depende únicamente de una autenticación previa, sino de la comprobación efectiva de las condiciones definidas por el modelo de gobernanza del espacio de datos.

Además, el framework se alinea con la lógica del intercambio soberano de datos implementada mediante conectores de espacio de datos. Los conectores gestionan activos, políticas y definiciones contractuales, y facilitan la negociación y transferencia entre participantes. Sin embargo, EDS4Agro complementa esta capacidad incorporando una capa de gobernanza adicional que permite controlar de forma más fina el acceso real a recursos concretos. Esta decisión arquitectónica refuerza la capacidad del espacio de datos para adaptar sus mecanismos de acceso a escenarios más complejos y a requisitos de seguridad y gobernanza más exigentes, sin depender exclusivamente de las capacidades nativas del conector.

Desde el punto de vista de la gobernanza del espacio de datos, este framework permite trasladar reglas y políticas de alto nivel a controles aplicables en tiempo de ejecución. Es decir, actúa como el nexo entre el marco normativo y contractual del ecosistema y el comportamiento efectivo de los componentes técnicos. Gracias a ello, las condiciones de acceso no quedan limitadas a una declaración documental, sino que pueden ser evaluadas y ejecutadas de forma automatizada, contribuyendo a la coherencia entre el diseño de gobernanza y la operación real del espacio de datos.

Conectores de espacio de datos

Los conectores de espacio de datos constituyen el elemento clave que habilita el intercambio efectivo de información entre los participantes del ecosistema EDS4Agro. En esta arquitectura, se ha adoptado la tecnología **Eclipse Dataspace Connector (EDC)** como base para la implementación de los mecanismos de descubrimiento, negociación y transferencia de datos, alineándose con los principios de soberanía, seguridad y confianza propios de los espacios de datos. Además, el uso de EDC responde a la necesidad de disponer de un componente que no solo facilite la conectividad técnica entre sistemas, sino que además incorpore capacidades nativas para la aplicación de políticas de acceso, la verificación de credenciales y la gestión de

relaciones de confianza entre participantes. En este sentido, el conector no actúa como un simple canal de comunicación, sino como el elemento que coordina el intercambio de datos bajo un modelo de autorización basado en credenciales verificables y políticas definidas por los proveedores de datos.

Desde el punto de vista de estándares y alineamiento arquitectónico, el EDC se fundamenta en las directrices del modelo **IDS-RAM (International Data Spaces Reference Architecture Model)**, lo que garantiza que la solución implementada cumple con los principios de interoperabilidad, seguridad y soberanía establecidos en el ecosistema europeo de espacios de datos.

Así, el conector proporciona una serie de funcionalidades esenciales para la operación del espacio de datos, entre las que destacan:

- La definición y exposición de un catálogo de datos propio por parte de cada participante.
- La descubierta de catálogos de otros participantes.
- La negociación de condiciones de acceso a los activos.
- La transferencia de datos entre proveedores y consumidores.
- La verificación de credenciales y aplicación de políticas de acceso.

Estas capacidades permiten establecer un entorno descentralizado en el que cada participante mantiene el control sobre sus datos, configurando de forma autónoma las condiciones bajo las cuales estos pueden ser compartidos.

Además, el conector se integra con el componente IdentityHub, responsable de gestionar las credenciales verificables (VCs) y generar las presentaciones verificables (VPs) necesarias para demostrar el cumplimiento de las políticas de acceso durante las interacciones entre participantes. Esta integración permite que las decisiones de acceso no dependan únicamente de mecanismos de autenticación, sino de un modelo de autorización verificable basado en identidad descentralizada.

Por otro lado, uno de los elementos fundamentales del funcionamiento del conector es la gestión del catálogo de datos, que se construye a partir de tres conceptos principales: activos (assets), políticas (policies) y definiciones de contrato (contract definitions).

Los activos representan los datos o servicios que un participante desea compartir dentro del espacio de datos. No se corresponden directamente con el dato en sí, sino con una representación abstracta que describe el recurso disponible, incluyendo metadatos relevantes que facilitan su descubrimiento y uso.

Las políticas definen las condiciones bajo las cuales un activo puede ser accedido o utilizado. Estas políticas establecen restricciones relacionadas con quién puede acceder al dato, con qué propósito y bajo qué requisitos de identidad o atributos. En este contexto, las políticas están estrechamente vinculadas con el modelo de identidad y confianza del espacio de datos, ya que su evaluación se basa en la información contenida en las credenciales verificables presentadas por los participantes.

Por su parte, las definiciones de contrato actúan como el elemento que vincula los activos con las políticas. A través de estas definiciones se establece formalmente qué activos pueden ser ofrecidos bajo determinadas condiciones, configurando así las entradas del catálogo que estarán disponibles para su descubrimiento y posterior negociación.

Pasando a lo relativo a la negociación y transferencia de datos, cabe destacar que el proceso de intercambio de datos en EDS4Agro sigue un ciclo de vida estructurado que comienza con el descubrimiento del catálogo y culmina con la transferencia efectiva de los datos.

En primer lugar, un participante consumidor realiza una consulta de catálogo a otro participante proveedor, identificando los activos disponibles y sus condiciones de acceso. Este proceso se realiza mediante los protocolos del dataspace implementados por el conector, permitiendo una interacción estandarizada entre nodos.

Una vez identificado el activo de interés, se inicia una fase de negociación de contrato, en la que se validan las condiciones de acceso definidas por el proveedor. Durante esta fase, el consumidor debe demostrar que cumple con las políticas establecidas, lo que implica la presentación de Verifiable Presentations (VPs) generadas a partir de sus credenciales verificables.

El proveedor, a través de su conector, verifica dichas presentaciones, comprobando tanto su validez criptográfica como su adecuación a las políticas definidas. Este proceso permite garantizar que el acceso al dato se concede únicamente a entidades autorizadas y bajo las condiciones previamente establecidas.

Una vez finalizada la negociación con resultado positivo, se inicia la transferencia de datos, que se realiza a través de los endpoints definidos en el conector del proveedor. En el caso de EDS4Agro, esta transferencia puede implicar el acceso a recursos gestionados por el Context Broker u otros sistemas internos del participante, garantizando que el consumidor recibe la información de forma segura y conforme a las condiciones acordadas.

Gestión del dato y contexto

La gestión del dato en EDS4Agro se basa en un enfoque orientado a contexto que permite representar, almacenar y compartir información de forma interoperable tanto a nivel técnico como semántico. Este enfoque se articula en torno al uso de un Context Broker NGSI-LD, un modelo de datos estructurado basado en grafos de conocimiento y el uso de vocabularios semánticos normalizados.

El núcleo de la gestión del dato en el espacio de datos es el Context Broker Orion-LD, que actúa como el componente central encargado de gestionar el estado de las entidades del sistema en tiempo real. Esta tecnología, basada en el estándar ETSI NGSI-LD, permite operar con datos enlazados (linked data), facilitando la interoperabilidad y la trazabilidad de la información a lo largo de todo el ecosistema.

Así, el Context Broker desempeña varias funciones clave dentro de la arquitectura:

- Gestión del ciclo de vida de las entidades, permitiendo la creación, actualización y consulta de datos conforme a estructuras definidas.
- Validación estructural de la información, asegurando que las entidades cumplen con el formato NGSI-LD y JSON-LD.
- Punto de acceso unificado a los datos, centralizando la interacción entre los distintos componentes del sistema y garantizando coherencia en el acceso a la información.

Este componente se apoya en un sistema de almacenamiento persistente basado en bases de datos NoSQL (MongoDB), lo que permite gestionar grandes volúmenes de datos agrícolas con estructuras flexibles y escalables.

Por otro lado, respecto al modelado, la representación de los datos se realiza mediante el el estándar NGSI-LD, utilizando el formato JSON-LD como mecanismo de serialización. Este modelo permite estructurar la información como un grafo de conocimiento, en el que las entidades se relacionan entre sí mediante identificadores únicos (URIs/URNs). Cada entidad del sistema se define mediante:

- Un identificador único (por ejemplo, urn:ngsi-ld:<tipo>:<UUID>).
- Un conjunto de propiedades, que representan atributos (estáticos o dinámicos).
- Un conjunto de relaciones, que enlazan la entidad con otras dentro del grafo.

Este enfoque permite modelar de forma flexible y extensible elementos del dominio agroalimentario, como parcelas, cultivos, sensores o predicciones, facilitando la trazabilidad completa desde la captura del dato en campo hasta su consumo final.

Además, el uso de JSON-LD permite incorporar contexto semántico directamente en la estructura del dato, facilitando su interpretación automática por distintos sistemas y garantizando la interoperabilidad entre participantes del espacio de datos.

Por otro parte, más allá de la interoperabilidad técnica, EDS4Agro incorpora mecanismos para garantizar la interoperabilidad semántica, es decir, que los datos compartidos tengan un significado unívoco y común para todos los participantes.

Para ello, se ha integrado el vocabulario AGROVOC, el tesauro de la FAO ampliamente reconocido en el ámbito agroalimentario. Este vocabulario se utiliza para enriquecer semánticamente los datos, vinculando conceptos y atributos a identificadores normalizados (URIs). En la práctica, esto implica que:

- Los términos utilizados en los modelos de datos pueden asociarse a conceptos definidos en AGROVOC.
- Cuando un concepto no está cubierto por modelos estándar, se enlaza directamente a su correspondiente URI en el tesauro.
- Se garantiza que diferentes participantes interpretan de forma homogénea los mismos conceptos (por ejemplo, tipos de cultivo, variables agronómicas o prácticas agrícolas).

De esta forma, NGSI-LD define la estructura de los datos, mientras que AGROVOC define su significado, permitiendo construir un ecosistema en el que los datos no solo pueden intercambiarse, sino también ser comprendidos y reutilizados de forma consistente.

Flujo de intercambio de datos

El intercambio de datos en el espacio de datos EDS4Agro se articula a través de un proceso estructurado que permite a los participantes compartir información de forma controlada, segura y conforme a las condiciones definidas por los proveedores. Este proceso se basa en las capacidades proporcionadas por los conectores de espacio de datos y sigue un ciclo de vida bien definido que abarca desde la publicación de activos hasta la transferencia efectiva de los datos.

En primer lugar, el proceso se inicia con la **publicación de activos de datos** por parte del participante proveedor. En esta fase, el proveedor define los recursos que desea compartir dentro del espacio de datos, representándolos como activos (assets) dentro de su conector. Estos activos no contienen necesariamente el

dato en sí, sino una descripción del recurso disponible, junto con la información necesaria para su acceso. A cada activo se le asocian políticas de uso y definiciones de contrato, que establecen las condiciones bajo las cuales podrá ser consumido por otros participantes.

Una vez publicados los activos, estos pasan a formar parte del **catálogo del proveedor**, que puede ser consultado por otros participantes del espacio de datos. A continuación, tiene lugar la fase de **descubrimiento**, en la que un participante consumidor realiza consultas sobre los catálogos disponibles con el objetivo de identificar activos de interés. Este proceso se realiza mediante los mecanismos de catálogo implementados por los conectores, permitiendo acceder a la información descriptiva de los activos y a sus condiciones de uso sin necesidad de acceder directamente al dato.

Tras la identificación de un activo relevante, se inicia la fase de **negociación**, en la que el consumidor solicita el acceso al recurso conforme a las condiciones definidas por el proveedor. Durante esta fase, el consumidor debe demostrar que cumple con los requisitos establecidos en las políticas asociadas al activo. Esto implica, en el modelo implementado en EDS4Agro, la presentación de credenciales verificables y la generación de presentaciones verificables que acrediten los atributos necesarios para acceder al dato.

El conector del proveedor evalúa esta información, verificando tanto la validez de las credenciales como su adecuación a las políticas definidas. En caso de que se cumplan las condiciones, se procede a la formalización del acceso mediante la creación de un **acuerdo o contrato de datos**, que recoge las condiciones específicas bajo las cuales se autoriza el acceso al recurso. Este contrato actúa como elemento vinculante entre proveedor y consumidor, estableciendo las reglas de uso que deberán respetarse durante la interacción.

Una vez formalizado el contrato, se habilita la fase de **transferencia de datos**, en la que el consumidor puede acceder al recurso conforme a las condiciones acordadas. La transferencia se realiza a través de los endpoints definidos por el proveedor, pudiendo implicar tanto la descarga de datos como el acceso a servicios o consultas sobre sistemas internos, como el Context Broker. En todo momento, el acceso se encuentra sujeto a las políticas y restricciones definidas, garantizando que el uso del dato se ajusta a lo establecido en el contrato.

Identidad, seguridad y control de acceso

La arquitectura del espacio de datos EDS4Agro incorpora un conjunto de mecanismos integrados que permiten garantizar la seguridad de las interacciones, el control efectivo del acceso a los recursos y la protección de la información intercambiada entre participantes. Estos mecanismos se basan en la combinación de autenticación robusta, autorización basada en políticas y verificación de credenciales, configurando un modelo de acceso coherente con los principios de soberanía del dato.

Desde el punto de vista de la autenticación, el sistema se apoya en la utilización de servicios de identidad que permiten verificar la identidad de los participantes antes de cualquier interacción. Este proceso se realiza mediante el uso de tokens de acceso emitidos por el proveedor de identidad, que son utilizados por los distintos componentes del sistema para validar la identidad del solicitante en cada petición.

Sin embargo, la autenticación por sí sola no es suficiente para garantizar el acceso a los recursos. Por ello, la arquitectura incorpora mecanismos de autorización basados en políticas, que permiten evaluar si un participante autenticado cumple con las condiciones necesarias para acceder a un determinado activo. Estas decisiones se apoyan en la evaluación de atributos del participante, del recurso y de la acción solicitada, permitiendo implementar un control de acceso de grano fino alineado con el modelo ABAC descrito en la arquitectura.

En este contexto, las decisiones de acceso se basan en la información aportada por los participantes a través de credenciales verificables y en los atributos contenidos en los tokens de autenticación. Esto permite que el sistema valide no solo la identidad del solicitante, sino también sus capacidades, roles o características relevantes para el acceso al recurso, garantizando que las políticas definidas por los proveedores se aplican de forma efectiva.

El modelo de seguridad se extiende también al propio proceso de intercambio de datos, en el que los conectores juegan un papel fundamental. Durante la negociación y transferencia, los conectores verifican las condiciones de acceso y aseguran que únicamente se permite el intercambio cuando se han cumplido las políticas establecidas. Este control se mantiene durante todo el ciclo de vida de la interacción, evitando accesos no autorizados o usos indebidos de la información.

Adicionalmente, la arquitectura contempla la protección de las comunicaciones entre componentes mediante el uso de protocolos seguros, garantizando la confidencialidad e integridad de los datos en tránsito. Este aspecto es especialmente relevante en un entorno distribuido, donde las interacciones se producen entre sistemas pertenecientes a distintas entidades.

Por otro lado, el modelo de control de acceso se integra con los mecanismos de gestión del dato y con los sistemas internos de los participantes, permitiendo aplicar restricciones no solo a nivel de acceso al recurso, sino también sobre las operaciones que pueden realizarse sobre el mismo. Esto permite adaptar el nivel de control a las necesidades específicas de cada activo y a las condiciones definidas en los contratos de datos.

Trazabilidad y monitorización

La arquitectura del espacio de datos EDS4Agro incorpora una capa específica de trazabilidad y monitorización orientada a registrar de forma verificable el uso de los activos compartidos y las interacciones entre participantes. Esta capacidad se implementa mediante una infraestructura de auditoría distribuida basada en tecnologías de registro distribuido (DLT), concretamente sobre Hyperledger Fabric, que permite garantizar la inmutabilidad, integridad y auditabilidad de los registros generados durante el intercambio de datos.

Esta capa de trazabilidad se concibe como un componente transversal de la arquitectura, complementando los mecanismos de identidad, negociación contractual y control de acceso. Su función no es almacenar los datos agrícolas transferidos ni sustituir a los sistemas operativos de gestión del dato, sino registrar las evidencias técnicas y contractuales asociadas a cada operación de acceso a los recursos del espacio de datos. De este modo, se establece una separación clara entre el plano de datos y el plano de auditoría, permitiendo reconstruir posteriormente el contexto completo de una interacción.

Desde el punto de vista arquitectónico, la solución se basa en una red permissioned de Hyperledger Fabric, en la que los nodos participantes están identificados mediante certificados digitales gestionados por el correspondiente Membership Service Provider. Sobre esta infraestructura se despliega un smart contract de auditoría, encargado de registrar y consultar las transacciones asociadas al intercambio de datos. Para facilitar la integración con el resto de componentes del sistema, se incorpora además una capa intermedia basada en una API REST, que actúa como pasarela entre los servicios del espacio de datos y la red blockchain, desacoplando la lógica de negocio del acceso directo al ledger.

El almacenamiento de las transacciones se realiza utilizando CouchDB como base de datos de estado, lo que permite conservar los registros en formato estructurado (JSON) y habilitar consultas avanzadas sobre los mismos. Para optimizar el acceso a la información, se definen índices sobre los principales campos del modelo de datos, facilitando la ejecución de búsquedas por participante, activo, acuerdo contractual o intervalo temporal.

Desde un punto de vista operativo, el registro de eventos se integra con el flujo de intercambio de datos del espacio de datos. Cada vez que un consumidor descubre un activo, negocia un contrato, obtiene autorización y realiza una transferencia efectiva, se genera un evento de auditoría que es enviado a la capa de trazabilidad y registrado en la red blockchain. Este mecanismo permite asociar cada acceso a un activo con un registro inmutable que contiene tanto información técnica de la operación como información contractual y de autorización derivada del flujo de los conectores.

El modelo de datos de las transacciones registradas permite capturar de forma completa el contexto de cada operación. Cada registro incluye, entre otros elementos, identificadores únicos de la transacción, marcas temporales, información sobre el recurso accedido, identificadores de consumidor y proveedor (DID), datos de red, identificadores de autenticación, referencias a los procesos de transferencia, identificadores contractuales, información sobre el activo, tipo de flujo de transferencia, metadatos del contenido y el resultado de la operación (código de estado y posibles errores).

Además de su función de registro, la infraestructura proporciona capacidades de consulta y análisis de las transacciones almacenadas. A través de los endpoints disponibles, es posible recuperar registros individuales, obtener el historial completo de interacciones, filtrar transacciones por distintos criterios (consumidor, proveedor, activo, acuerdo, etc.) o realizar consultas por rangos temporales. Asimismo, se dispone de funcionalidades para la obtención de métricas agregadas del sistema, como el número total de transacciones, el volumen de datos transferidos o el número de participantes activos.