IMAGEN DE BENZOIX EN FREEPIK

DOCUMENTO

ESPACIOS DE DATOS COMPARTIDOS
RETOS DE SEGURIDAD

ENERO DE 2024

Ametic
LA VOZ DE LA INDUSTRIA DIGITAL

Espacios de datos compartidos: retos de seguridad

Introducción

Se denominan espacios de datos compartidos a aquellos entornos técnicos en los cuales es posible **compartir datos, conservando la soberanía sobre ellos**, y en unas condiciones técnicas y normativas que refuercen la **confianza** en el intercambio de datos, con el fin de proporcionar nuevos servicios, y promover nuevos negocios. En teoría, los espacios de datos no se basan en ninguna infraestructura determinada, sino que aspiran a que se pueda partir de cualquier infraestructura previa, y que los módulos que se añadan, convenientemente probados y certificados, sean los que proporcionen la capacidad de interconexión, salvaguardando la interoperabilidad semántica, así como haciendo visibles los datos a los posibles utilizadores y consumidores, según las condiciones estipuladas por quien ostente la soberanía.

Los espacios de datos deben garantizar el intercambio de información entre proveedores y consumidores de datos de manera segura, porque es la clave a la hora de garantizar la confianza, igual que es la salvaguarda de la privacidad. Así como las organizaciones deben de asegurar la privacidad y seguridad de los datos que almacenan y gestionan, los espacios de datos compartidos deben proporcionar garantías de ciberseguridad que permitan y garanticen la privacidad y el intercambio de información de manera segura, habilitando a las partes determinar el cómo, cuándo y en qué condiciones se utilizan los datos.

El término **conector** procede del entorno de Industrial Data Spaces (IDS), un paradigma pionero en la formulación y definición de qué son espacios de datos compartidos. El mismo término se empieza también a utilizar en Gaia-X, que podría considerarse la expansión de los espacios de datos hacia la federación, con el fin de ampliar las posibilidades de utilización de datos hasta, hipotéticamente, todos los posibles. Tanto IDS, fomentado por la asociación IDSA, como Gaia-X, detrás de la cual está la asociación Gaia-X AISBL, de origen europeo, y respaldada por la Comisión, pero con alcance transfronterizo creciente, están en fase de consolidación, aunque se benefician del trabajo conjunto.

Espacios de datos según Industrial Data Spaces Association

Un requisito estratégico de los espacios de datos industriales (IDS) es proporcionar **cadenas de suministro de datos seguras**. Esto es fundamental para establecer y mantener la confianza entre los participantes que desean intercambiar y compartir datos y utilizar aplicaciones sobre esos datos. La arquitectura de seguridad de IDS proporciona medios para identificar dispositivos en el entorno de IDS, proteger las transacciones de comunicación e intercambio de datos, y controlar el uso de los datos después de que se hayan intercambiado.

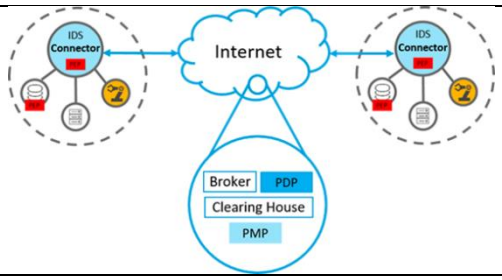
Los conectores de IDS aseguran que especificaciones y requerimientos de arquitectura de seguridad se materialicen en interacciones y operaciones cotidianas en espacios de datos. Esta es la función de conector de IDS en la arquitectura (Figura 1):	

Figura 1: Función del conector (Fuente: IDSA)

La seguridad es crucial, si bien **no se proporcionan servicios específicos o herramientas de seguridad o ciberseguridad** en general, sino que estas deben ser consideradas al diseñar y desarrollar el software. Dado que la seguridad generalmente cubre aspectos no funcionales, la seguridad y la confianza sirven como habilitadores de funcionalidades como el intercambio de datos. En consecuencia, la seguridad se trata como un requisito transversal de las capas del modelo de arquitectura de referencia (IDS-RAM).

Los aspectos de seguridad y el modelo de confianza se usan para permitir o restringir algunas transacciones u operaciones. Sin esta seguridad, muchos casos de uso no serían posibles, como ofrecer **datos confidenciales a socios comerciales confiables**. Por ejemplo, para permitir la identificación y autenticación confiables de los componentes mediante una infraestructura de clave pública (PKI), el operador de este componente debe generar un par de claves en el componente, solicitar un certificado de clave pública de la Autoridad de certificación (CA) y proporcionar este certificado en el componente. Para el soporte de atributos dinámicos, el proveedor del servicio de aprovisionamiento de atributos dinámicos (DAPS) debe verificar los atributos que confirmará con los tokens de atributos dinámicos (DAT). Lo mismo se aplica a las operaciones confiables de una tienda de aplicaciones, para las cuales una entidad confiable debe verificar y firmar los datos antes de que se puedan cargar. Y así sucesivamente.

De hecho, en el proceso de certificación de los conectores hay un paso de autoevaluación (*self assessment*) que consiste básicamente en una lista de **verificación de las especificaciones de seguridad identificadas en DIN SPEC 27070**, como paso previo a la certificación de los IDS conectores desarrollados. DIN SPEC 27070 especifica los requerimientos y la arquitectura de referencia del *security gateway* para el intercambio de datos y servicios industriales, siguiendo la familia de estándares de seguridad de IEC62443-4-2

Evolución de los espacios de datos hacia su federación

Sobre la base de la creación de espacios de datos individuales, sectoriales, o de cadena de valor, se pretende que estos actúen de manera federada con el fin de ampliar las posibilidades de uso de datos. Para tal fin se creó el modelo Gaia-X, desarrollado por la asociación Gaia-X AISBL. Trabajando en el plano técnico junto con IDSA, FIWARE Foundation, y Big Data Value Association BDVA, dentro de la llamada Data Spaces Business Alliance, ha propuesto la arquitectura de alto nivel de Ecosistema de Mercado Abierto Descentralizado (DOME):

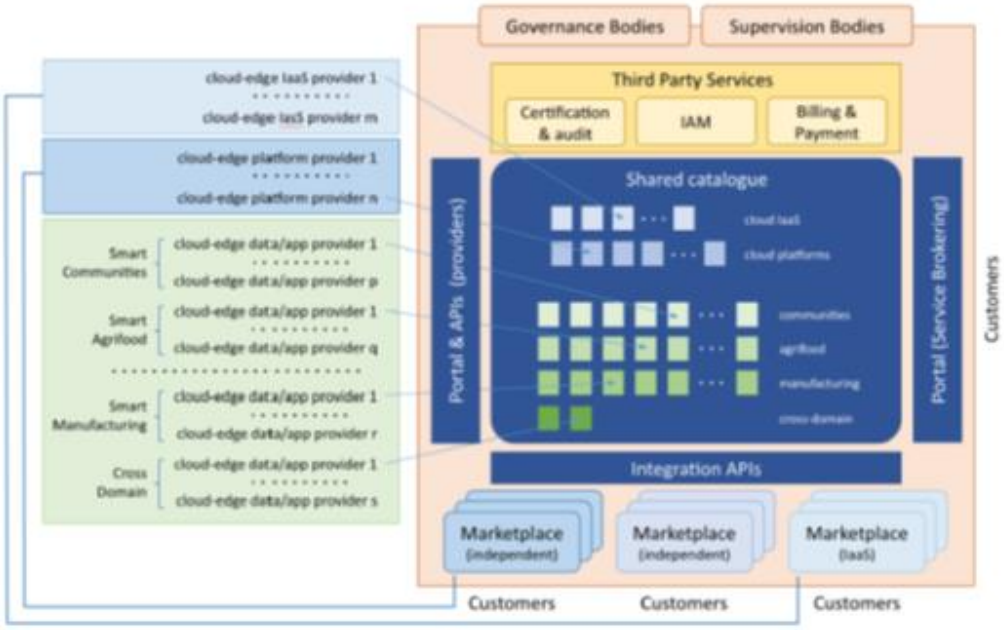


Figura 2: Decentralized Open Marketplace Ecosystem DOME (Fuente: DSBA)

Cualquier espacio de datos requiere un Marco de Anclaje de Confianza (Trust Anchor Framework) y su asociado Marco de Gestión de Identidad y Acceso Descentralizado, para permitir la operación confiable del sistema **sin requerir una entidad central intermediadora** de todas las interacciones entre los participantes. Esto garantiza la confianza en la información publicada en los espacios de datos por parte de los proveedores, así como que los clientes accedan a los servicios de portal de espacio de datos, administren su perfil, e iniciar sesión en los mercados federados de datos.

Dicho MAC (TAF) define y obliga a un conjunto de reglas que las diferentes organizaciones acuerdan seguir para brindar sus servicios, e incluye legislación, estándares, y orientación, entre otras reglas. Al seguirlas, todos los servicios y organizaciones participantes pueden usar sus identidades y atributos digitales de manera consistente y confiable. En este sentido, se está dedicando mucho esfuerzo al plano de confianza y, en consonancia, en los ***self-descriptors* para los participantes y los proveedores** de servicios.

Así pues, las características del servicio de conformidad de Gaia-X se resumen:

- Verificar el cuerpo de autodescripción de un participante o servicio de acuerdo con el marco de confianza.
- Verificar las firmas de autodescripción y las credenciales provenientes de *trust anchors*.
- Proporcionar una autodescripción canonizada y normalizada para la prueba de conformidad.
- Añadir una prueba de cumplimiento y emisión de Credencial Verificable (VC).
- Verificar de VC (autodescripción firmada, incluida la prueba de cumplimiento).
- El servicio se puede utilizar a través de API y Swagger UI.

Expectativas

El entorno al que se quiere llegar por medio de Gaia-X se representa en la Figura 3, y ofrece un marco de confianza con las siguientes características:

- Mensurable y comparable: índice de confianza de Gaia-X en función de la composición del servicio y la identidad de la firma.
- No toma decisiones por el usuario.
- Hace cumplir la transparencia.
- Habilita la portabilidad: composición del servicio.
- Habilita la trazabilidad: agregación de consentimiento y atenuación de políticas.
- Automatizable - lectura mecánica
- Seguro: principio de prueba de conocimiento cero
- Escalable: por semántica de tipo web web
- Rentable: inclusive para PYMEs
- Aporta subsanación y sanciones

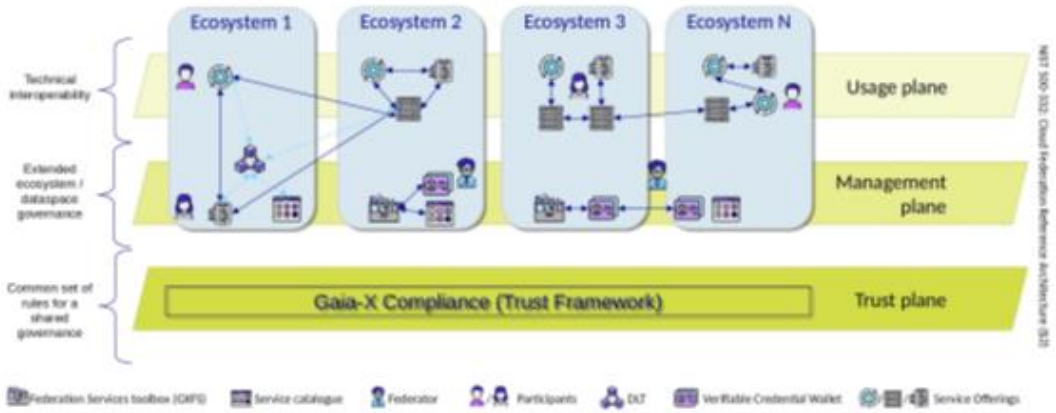


Figura 3: Escenario de compartición de datos en espacios federados (Fuente: Gaia-X ASIBL)

Carencias de desarrollo

En el desarrollo del Marco de Anclaje de Confianza TAF han surgido diferentes cuestiones, a las que se está tratando de dar soluciones. Como, por ejemplo:

- Vinculación de ID: ¿Cómo verificar que un identificador dado corresponde a una identidad legal válida de una entidad en el mundo real?
- Prueba de participación: ¿Cómo verificar que la entidad es confiable porque es un participante suscrito en un ecosistema determinado (por ejemplo, para verificar la confianza del Catálogo Compartido de Especificaciones de Productos y de Ofertas de Productos)?
- Prueba de Autoridad Emisora: ¿Cómo comprobar que las credenciales presentadas por un participante hayan sido emitidas por otra entidad que pueda ser considerada Emisora de Confianza de ese tipo de credenciales? Esto permite que el verificador deposite la confianza adecuada en los hechos atestiguados por las Credenciales Verificables presentadas por un participante.

En cuanto a la administración de identidad y acceso descentralizado hay que abordar:

- Identificación: ¿Cómo verificar que un identificador enviado por un participante a otra entidad ha sido enviado por el participante y no por un impostor que conoce el identificador? Además, necesitamos vincular criptográficamente el identificador a las Credenciales Verificables enviadas por el participante para que los hechos atestiguados en las credenciales puedan usarse para autenticación y autorización.
- Autorización: ¿Cómo usar los hechos certificados en las Credenciales verificables presentadas por un participante para realizar un control de acceso avanzado de RBAC/ABAC y aplicación de políticas?

Se ha propuesto confiar en los identificadores ya utilizados en los certificados digitales emitidos por los Proveedores de Servicios de Confianza (TSP) autorizados por las leyes europeas pertinentes. La combinación de certificados digitales emitidos por TSP y credenciales verificables contribuye a la validez legal y la interoperabilidad de las transacciones transfronterizas relacionadas con datos en la Unión Europea, lo que facilita la validación transfronteriza de firmas electrónicas, sellos electrónicos y más. Esencialmente, las credenciales y presentaciones verificables (incluidas las especificaciones y ofertas de productos) utilizadas en el ecosistema se firmarán mediante certificados digitales.

Tanto las especificaciones del producto como las descripciones de la oferta del producto podrán contar con una serie de etiquetas emitidas por las agencias de certificación en relación con el servicio ofrecido que certifica el cumplimiento de los reglamentos de la UE definidos o las reglas establecidas por las autoridades de supervisión (RGPD, reglamentos de sectores específicos, Cloud Rulebook de la UE, …), estándares relevantes (por ejemplo, estándares de interoperabilidad), o mejores prácticas (Open Source Security Foundation Best Practices).

Dado que cualquier persona puede tener acceso a la tecnología necesaria para crear Credenciales Verificables y cualquiera puede emitir credenciales y firmarlas digitalmente con su certificado digital eIDAS, el problema es cómo un verificador sabe que las Credenciales Verificables han sido emitidas por una entidad facultado para expedir ese tipo de credencial. El principal mecanismo para resolver este problema es el uso de Listas de Emisores de Confianza.

Un problema adicional relacionado con la Vinculación de ID cuando un participante envía en línea una Credencial Verificable a otro participante es asegurarse de que la credencial haya sido enviada por una entidad autorizada para hacerlo y no por un impostor, para lo que se usará un enfoque basado en utilizar criptografía de clave pública y aprovechando la confianza que ya brindan los certificados digitales utilizados en el enlace de identificación.

En principio, se podría utilizar el mismo enfoque para las personas físicas que para las personas jurídicas. Los estándares ETSI también cubren a las personas físicas y definen un "Identificador semántico de persona física". Sin embargo, las personas jurídicas son completamente diferentes a las personas físicas, especialmente desde el punto de vista de la privacidad (vide diferencias en RGPD). Por esas razones de privacidad se ha de proponer un enfoque diferente para los identificadores de personas físicas.

Otros inconvenientes surgen al verificar una credencial/presentación verificable, como:

1. ¿Cómo saber que el emisor de la Credencial Verificable es un participante en el ecosistema concreto (por ejemplo, un Espacio de Datos dado) donde también se es participante?

2. ¿Cómo saber que el sujeto de la credencial verificable es un participante en el ecosistema concreto (por ejemplo, un Espacio de Datos dado) donde también se es participante?

Para esto, se propone utilizar una Lista de participantes de confianza que incluya las identidades y los metadatos asociados de todas las personas jurídicas que participan en el ecosistema concreto.

Por último, cabe señalar que, actualmente, el conector se identifica mediante los atributos de un certificado X.509. El identificador basado en DID (identificador descentralizado, según esquema URI [*universal resource identifier*] de W3C) tiene que ser descrito. Esto todavía está abierto. También que, en IDS, el perfil de seguridad es validado por una agente de evaluación externa y proporcionado a una autoridad central, pero el flujo de trabajo de proporcionar la asignación directamente como VC/VP debe describirse en detalle.

Oportunidades de negocio y mensajes para los primeros receptores (Comisión Ciberseguridad)

Como se ha visto en el apartado anterior, hay una notable cantidad de aspectos por resolver, que se encuentran aún en manos de los grupos operativos de Gaia-X, y que están abiertos a la participación de los agentes que lo deseen.

Cuando esté consolidada, la arquitectura propuesta (DOME) proporcionará medios para la integración de servicios de terceros, que pueden resultar de interés para empresas del sector de ciberseguridad y conformidad, como, por ejemplo:

- Servicios de agencias de certificación y auditoría que ayuden a validar la confiabilidad, seguridad y soberanía de ciertos servicios en la nube mediante la comprobación y verificación de su cumplimiento con certificaciones predeterminadas en todo el mercado.
- Proveedores de servicios de IAM (*identity and access management*) que ofrezcan servicios alineados con estándares abiertos para IAM adoptados en DOME, brindando a los participantes la capacidad de administrar de manera segura identidades y acceder a servicios específicos de aplicaciones y datos en la nube y perimetrales.
- Proveedores de servicios de facturación y pago que funcionan como puertas de enlace que se basan en registros de transacciones registrados en la infraestructura de red de cadena de bloques federada subyacente a DOME, para proporcionar facturación segura, transparente y confiable a los consumidores y pago a los proveedores.

Para esos tres tipos de servicios de terceros, o adicionales, DOME representa una nueva fuente de ingresos, ya que les da acceso a un nuevo mercado (los proveedores de servicios en la nube y de *borde,* y los clientes). Por otro lado, pueden representar fuentes potenciales de ingresos para asegurar la sostenibilidad de DOME.

Documentación

"International-Data-Spaces-Association IDS-RAM 4.0 Security Perspective"

https://github.com/International-Data-Spaces-Association/IDS-RAM_4_0/blob/main/documentation/4_Perspectives_of_the_Reference_Architecture_Model/4_1_Security_Perspective/README.md

"Technical Convergence, Discussion Document, Version 1.0.1" Data Spaces Business Alliance, 2022-09-26

"Gaia-X Lab – Compliance Service" Gaia-X Lab, Gaia-X CTO Office, 2022-09-26