

EUNOMIA

Εὐνομία

Introducción a los espacios de datos federados y distribuidos : generando confianza digital e implementarlos con facilidad

Joaquín Salvachúa; Andres Muñoz, Gabriel Huecas
Universidad Politécnica de Madrid

Presentations

GING - UPM

Todos pertenecemos a este lugar: <https://ging.github.io/>

Grupo de investigación GING en la Universidad Politécnica de Madrid.

Nuestra investigación se centra principalmente en protocolos y estándares y tecnologías de WWW aplicados a numerosos casos de uso. Actualmente, nos centramos en la investigación en computación en la nube, educación, análisis del aprendizaje, ingeniería de datos, sistemas de videoconferencia distribuidos con estándares WWW, LLM e IA, datos abiertos y espacios de datos.

Involucrado en métodos formales de protocolo para la especificación, validación y verificación de protocolos que utilizan álgebras de procesos (LOTOS) hace algunos años. Participación en varios comités de normalización.

Participación en IETF, W3C, ETSI y otros organismos de normalización.tion,.



Presentations

NUESTRA MISIÓN

Llene los vacíos de investigación para hacer realidad los espacios de datos.

Desarrolle componentes que cualquiera pueda usar de la manera más fácil.

Integración con herramientas de Big Data y ML-OPS reconocidas.

Enfoque MPV en la implementación.

Project supported by INCIBE

EUNOMIA-Soluciones para la soberanía, confianza y seguridad en los espacios de datos

*C.128.23 EUNOMIA,
C130.23 MCIPYME*

Presentations

¿Qué es un espacio de datos?

Definición Wikipedia :

Un espacio de datos es un ecosistema donde materializar la compartición voluntaria de los datos de sus participantes dentro de un entorno de soberanía, confianza y seguridad, establecido mediante mecanismos integrados de gobernanza, organizativos, normativos y técnicos.

El espacio de datos es pieza clave para el desarrollo de la economía del dato, al habilitar su acceso, intercambio y reutilización legítima, posicionando al dato como recurso no rival, cuya utilidad crece según su uso se generaliza en un claro ejemplo de efecto red.

Definición en Gaia-X :

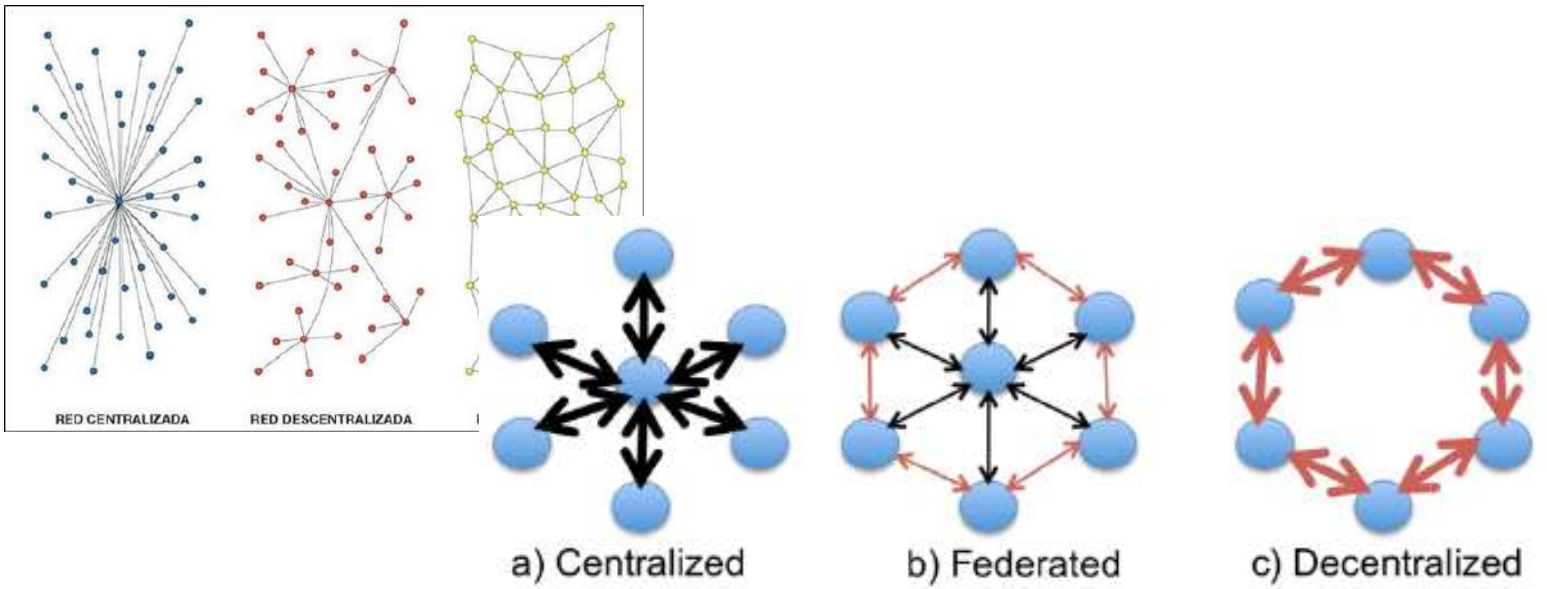
Permite una infraestructura de datos federada y segura, en la que los datos se comparten, y los usuarios conservan el control sobre el acceso y el uso de sus datos.

Permite la creación de vínculos entre muchos proveedores de servicios en la nube en un ecosistema más amplio, transparente y justo para impulsar la economía europea de los datos del mañana.

Presentations

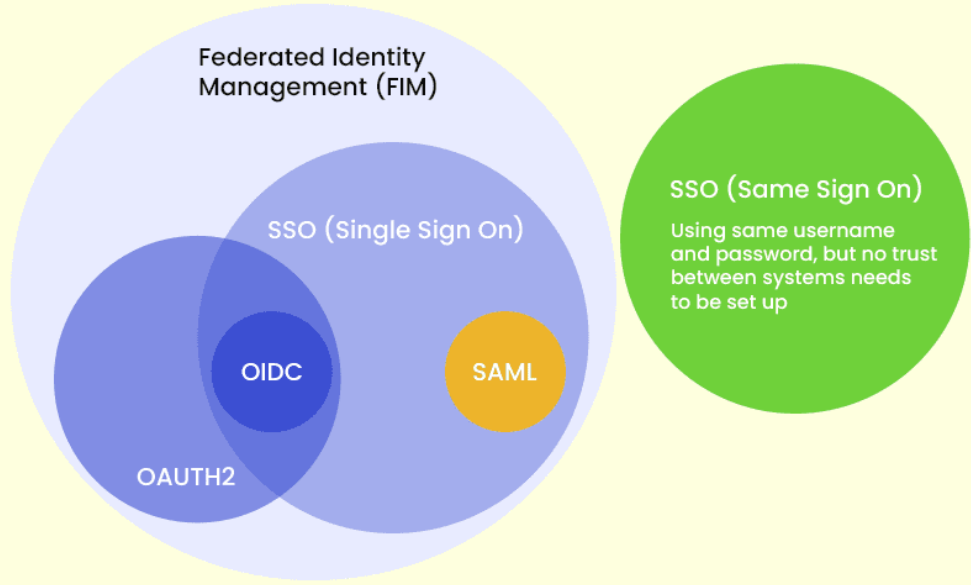
¿Qué es un espacio de datos?

Se utiliza a menudo para describir un espacio de datos como una infraestructura descentralizada para el intercambio y el intercambio de datos confiables en ecosistemas de datos basados en principios comúnmente acordados (Nagel y Lycklama 2021).



Presentations
Federación

SSO is a Part of Federated Identity Management Architecture



SAML - Security Assertion Markup Language

OIDC - OpenID Connect

OAuth2 - Open Authorization2

Presentations

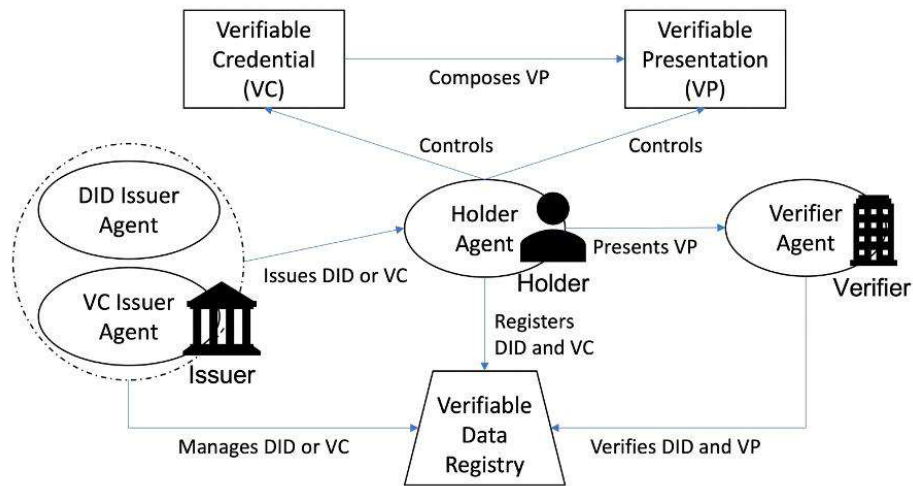
Modelo de presentación de credenciales. (W3C)



Modelo actual de Identidad

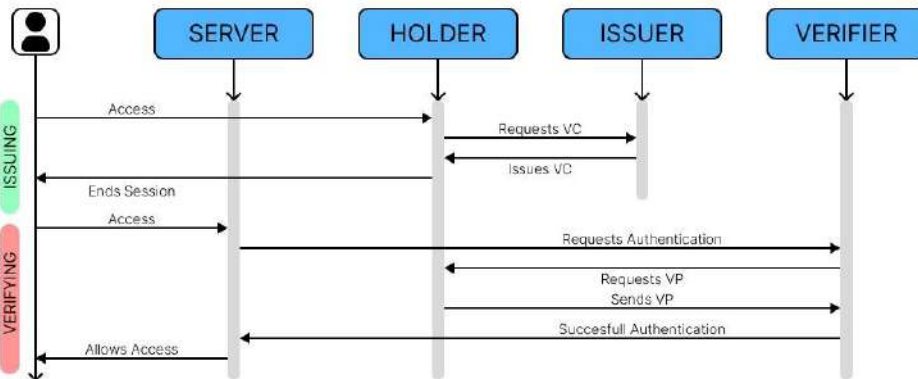


Modelo SSI (Auto-Identidad Soberana)



Desarrollo de nuevas propuestas del W3C

- Basado en W3C VC-API (definición en curso).
- NO se requiere SIOP / OID4VC ...
- Integración G NAP



ODRL. (W3C)

El Open Digital Rights Language (ODRL) es un lenguaje de expresión de políticas que proporciona un modelo de información, vocabulario y mecanismos de codificación flexibles e interoperables para representar declaraciones sobre el uso de contenido y servicios.

Inicialmente se centró en las políticas de consumo de medios a través de internet.

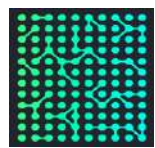
Ahora se están desarrollando varios perfiles para definir políticas en varios aspectos.

En curso una nueva semántica ODRL 3.0 con especificación temporal, entre otras capacidades más avanzadas.

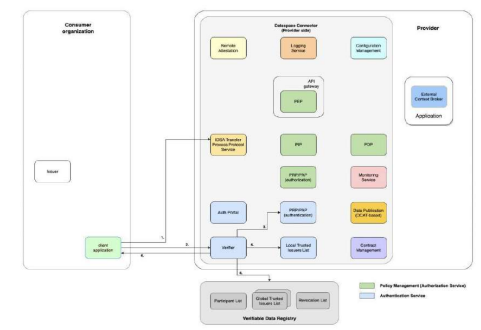
FIWARE, Data Spaces and Zero Trust

Pathway to provide ReBAC in FIWARE :

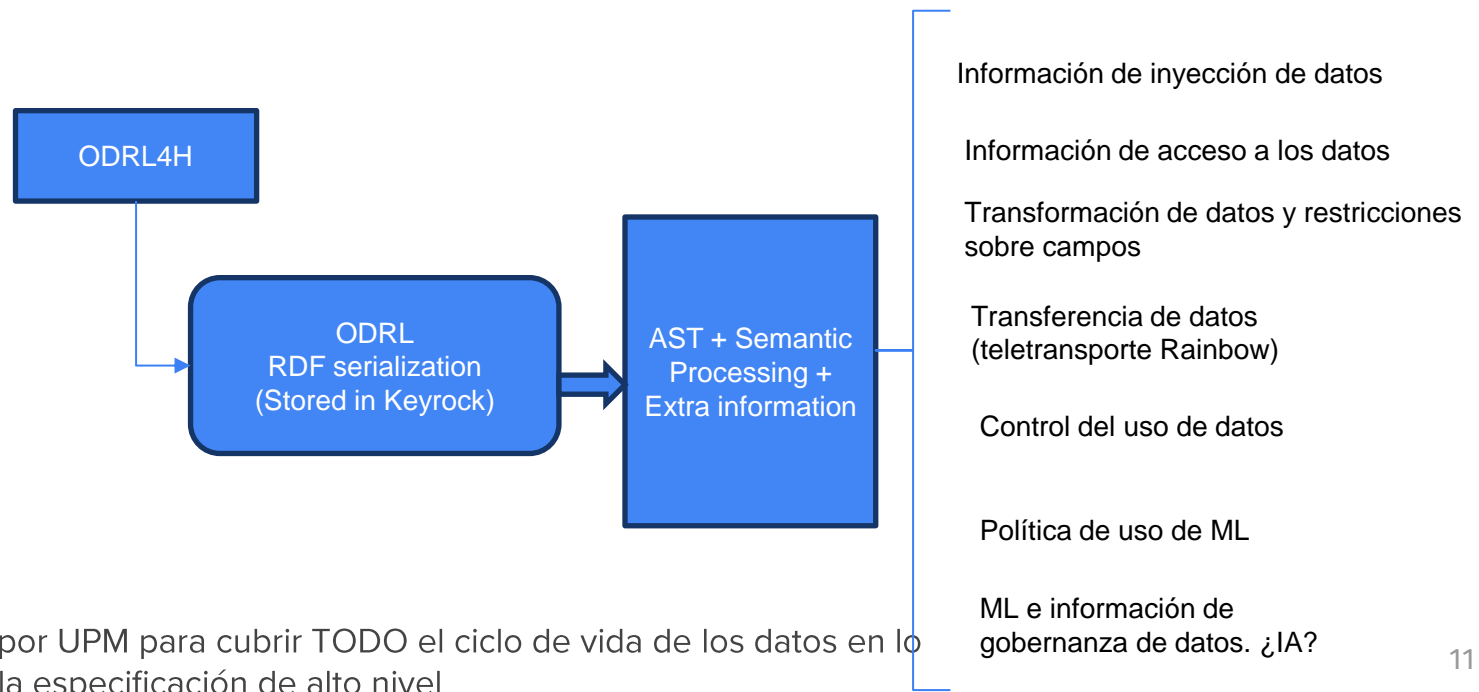
- Provide the logic of ReBAC in FIWARE Keyrock and Wilma Generic Enablers
- Integrate the authorization system of Open FGA in Keyrock
- Integrate Keyrock as authentication and authorization system in the FIWARE Data Space Connector
- Interconnection with SSI (more later).



Keyrock



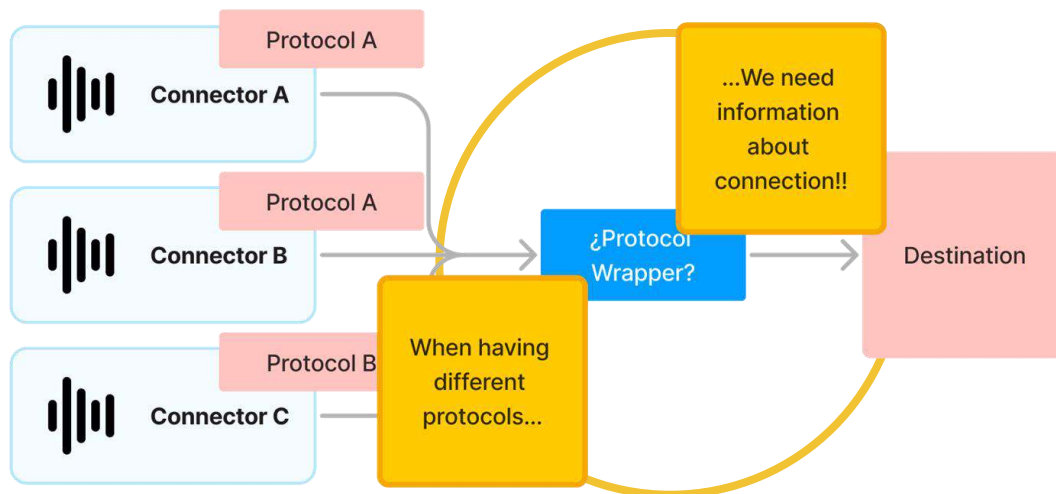
ODRL Para Espacios de Datos :



Perfil desarrollado por UPM para cubrir TODO el ciclo de vida de los datos en lo que respecta a la especificación de alto nivel

Transport protocol

Diferentes escenarios de transferencia



Necesidad de proporcionar detalles de los puntos de acceso y versiones de protocolos

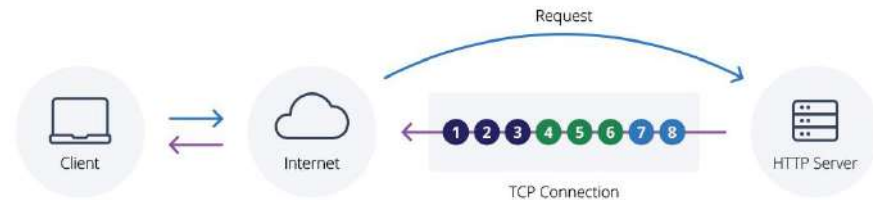
Ofrecer un traje de validación de pruebas basado en el banco de pruebas de las instalaciones de Interoperable Europe.

<https://joinup.ec.europa.eu/collection/interoperability-test-bed-repository/solution/interoperability-test-bed>

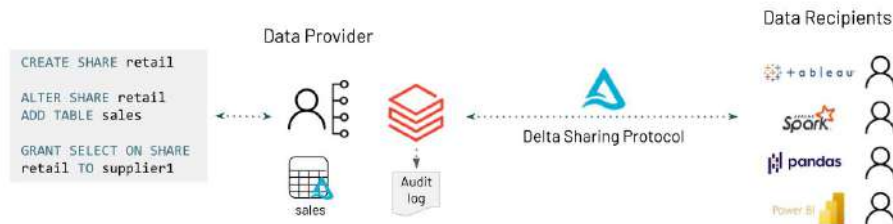
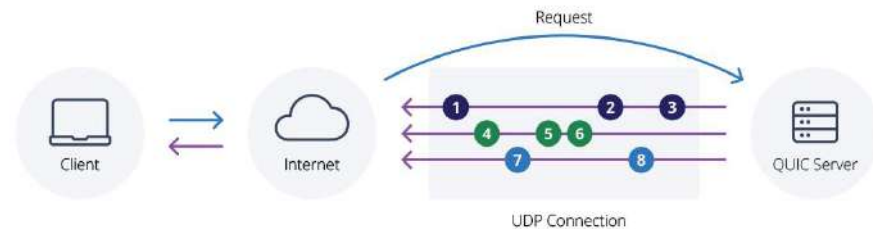
QUIC protocol (Quick UDP Internet Connections)

- Estándar IETF (RFC 9000) e implementado en la mayoría de los navegadores ahora.
- Nueva versión en curso (QUIC Versión 2 RFC 9369)
- Inicialmente desarrollado como reemplazo rápido de HTTP.
- No es necesario realizar modificaciones para nuestros fines
- También la integración en curso del protocolo Delta-Sharing

HTTP/2



QUIC



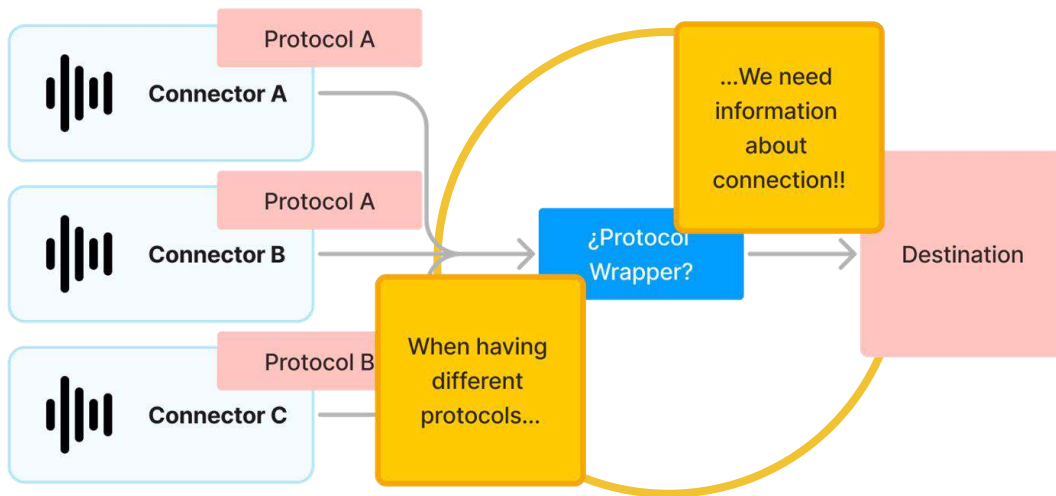
Protocolo de transporte de alta velocidad

Tarea de implementación

- Interconectado a la implementación del protocolo de espacio de datos
 - Implementación continua con Scala y Rust
 - Integración con conectores Apache Spark/Delta Lake: Proporciona una prueba de concepto para un ciclo de vida completo de ML-OPS.
 - Modificación de algunas implementaciones de servidores QUIC basados en RUST.
 - Integración en el conector de espacio de datos FIWARE.
-
- Entorno de despliegue SENCILLO de todo el conector.

Transport protocol

Different transfer scenarios



Need to provide **access points details** and **protocols versions**

Offer a test validation suit based on **Interoperable europe facilities test bed.**

<https://joinup.ec.europa.eu/collection/interoperability-test-bed-repository/solution/interoperability-test-bed>

Integración con herramientas de gobernanza

Necesidad de tener todos los eventos de gobernanza para una solución de ML completa.

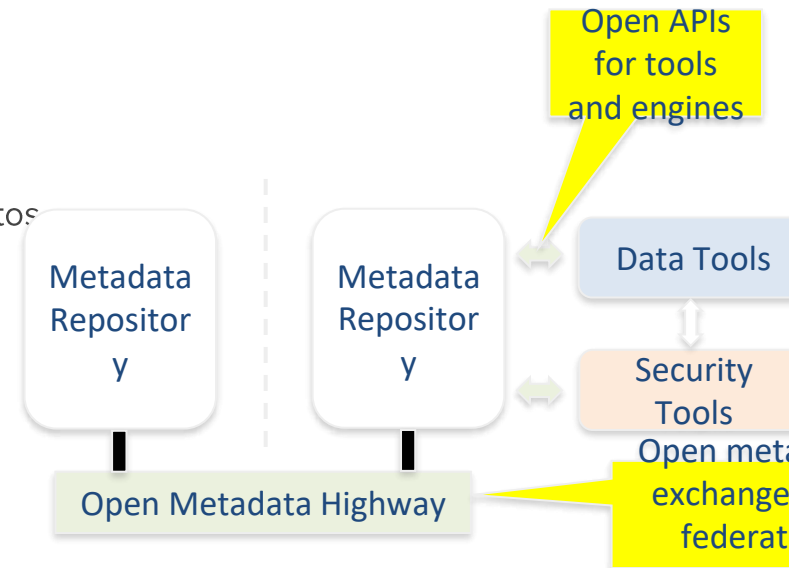
Se ha agregado un protocolo de transporte en el protocolo de espacio de datos para los metadatos de gobernanza y procedencia entre los puntos de conexión (persistentes en el tiempo).

Origen de la parte de control de uso de datos

Utilícelo como fuente de metadatos.

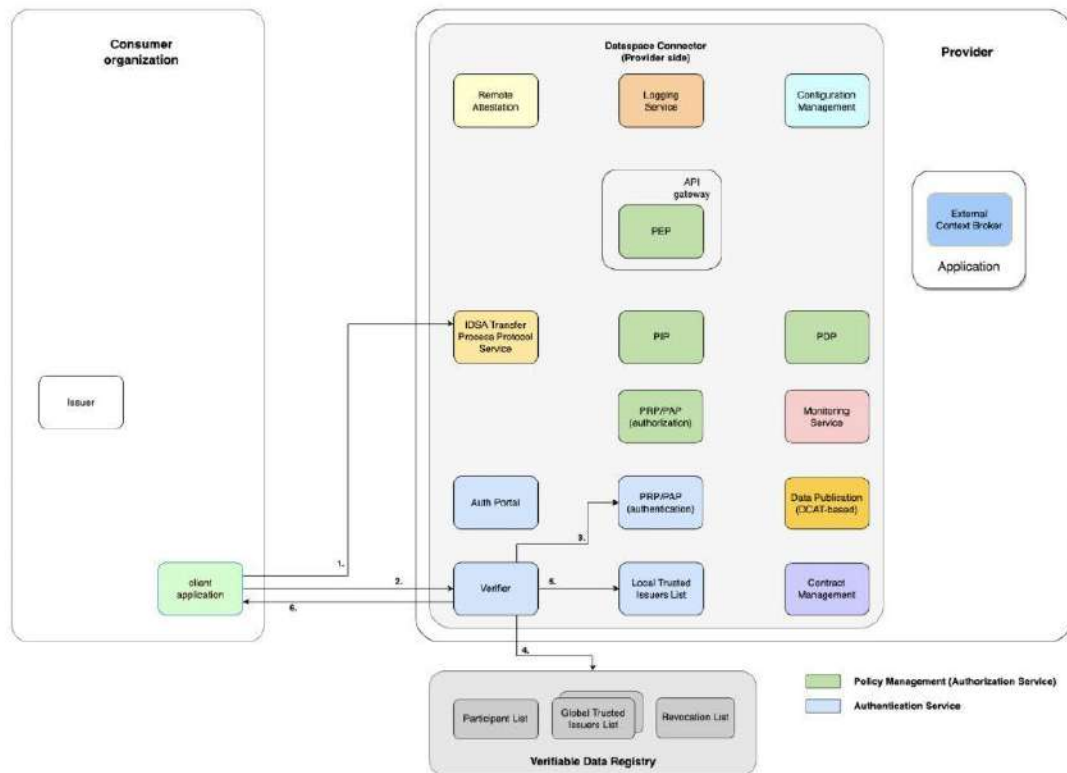
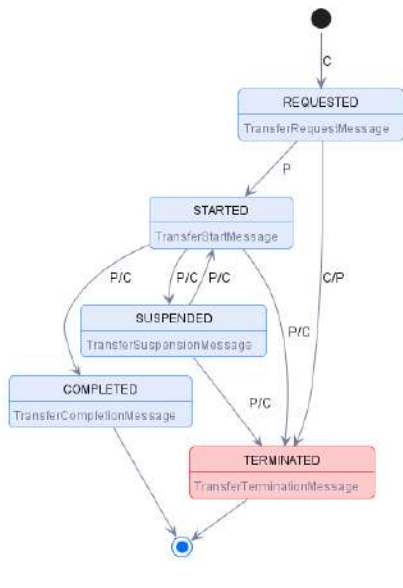
Considere la integración con otras herramientas de gobernanza de datos. Nuevo plano en el protocolo de transferencia.

Inclusión de información de linaje distribuido.



Transport protocol

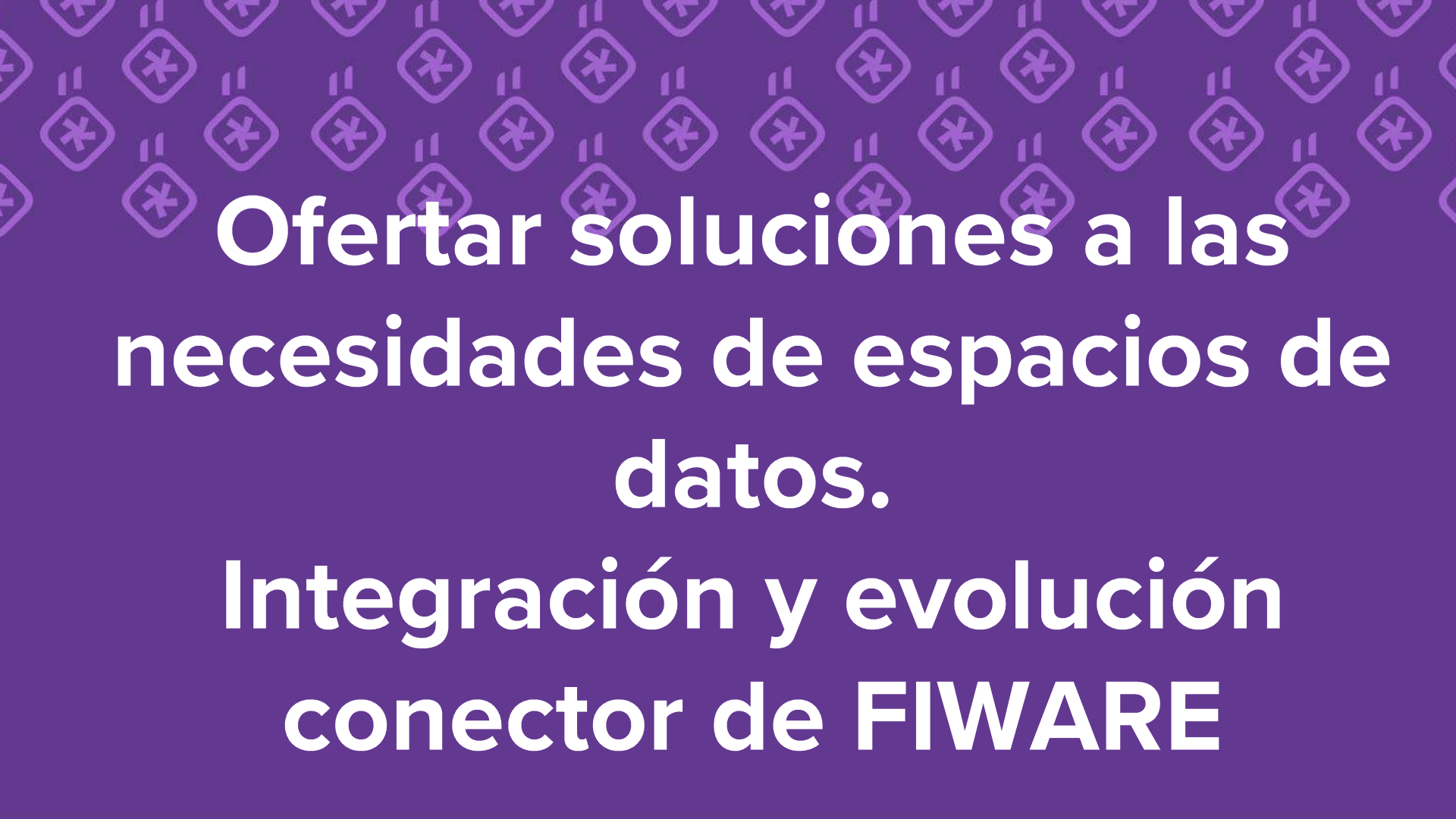
Integration of the solution FIWARE connector



Diferentes tareas

Trabajamos en varios niveles, intentando llenar algunos vacíos dentro de las arquitecturas del espacio de datos y predecir algunas líneas futuras, algunas de ellas como las siguientes:

- **Evolución de los protocolos de transporte**
- **Cobertura de los requisitos de gobernanza de datos y gobernanza de datos distribuidos**
- **Abordar los sistemas de anclaje de confianza más allá de las fronteras europeas y en el entorno de desarrollo y pruebas.**
- **Uso de la identidad autosoberana, SSI**
- **Aplicación de ODRL, Zero Trust y ReBAC para la gestión de políticas**
- **Metadatos para espacios de datos: fácil de generar**
- **Uso de arquitecturas DataLakehouse aplicadas a espacios de datos**
- **Sistemas de transmisión de alta velocidad**
- **Integraciones con conectores FIWARE**
- **Desarrollar un conjunto de pruebas para el protocolo de espacio de datos compatible con la Europa interoperable**



**Ofertar soluciones a las
necesidades de espacios de
datos.**

**Integración y evolución
conector de FIWARE**

Gracias 🙌

Preguntas en el café