

Presentación general Hércules ASIO

Esteban Sota

RIAM Intelearning Lab – GNOSS

estebansota@gnoSS.com



HĒRCULES



Hércules. Presentación general de Hércules ASIO

- ☐ Desarrollo de ASIO. UTE GNOSS y Universidad de Deusto.
- ☐ Proyecto Hércules.
- ☐ Hércules ASIO. Presentación.
- ☐ Datos abiertos y enlazados y Grafos de conocimiento.
- ☐ Hércules ASIO. Infraestructura Ontológica.
- ☐ Hércules ASIO. Arquitectura Semántica.



FONDO EUROPEO DE DESARROLLO REGIONAL (FEDER)

Una manera de hacer Europa

UTE GNOSS y Universidad de Deusto





- ☐ Somos una pyme tecnológica que lleva trabajando más de 10 años en la construcción de grafos de conocimiento y en la Web de los Datos (web Semántica)
- ☐ Que colabora con las organizaciones en transformar la inteligencia humana en inteligencia tecnológica o artificial...
- ☐ Con el objetivo de facilitar a las personas el acceso a una vida digital más feliz
- ☐ www.gnoss.com



Trabajamos en España, con
sedes en Bilbao y Logroño
(La Rioja)

gnos #VidaDigitalFeliz

GNOSS transforma el entorno con solidez, pensando que la tecnología es para las personas y no al revés, pensando en el largo plazo y con un objetivo claro: hacer fácil lo difícil, conseguir una **#VidaDigitalFeliz**



Transformación Digital
#VidaDigitalFeliz
[La Tecnología en su sitio]



Tecnología Fabricada en España.
MADE IN SPAIN
"SpanishTech Guerrilla"
Guerrilla Tecnológica Española
(Posicionamiento)



La IA habla idiomas
IA en Español
(Mercado)



¡EUREKA!
Todo está enlazado
Inteligencia artificial semántica
basada en la explotación de
Grafos de Conocimiento
(Visión tecnológica)

FONDO EUROPEO DE DESARROLLO REGIONAL (FEDER)

Una manera de hacer Europa



La tecnología no es neutral

Inteligencia Artificial para una experiencia digital feliz

FONDO EUROPEO DE DESARROLLO REGIONAL (FEDER)

Una manera de hacer Europa

gnos

CULTURA

MUSEO NACIONAL DEL PRADO

Plataforma digital y Web del Museo del Prado

PATRIMONIO NACIONAL

Patrimonio Nacional en la Web

MUSEO THYSSEN-BORNEMISZA

Modelo Ontológico de la Colección del Museo

Vivanco

La cultura del vino y sus contextos: colección y centro de documentación del Museo Vivanco enlazados en un grafo de conocimiento

ADMINISTRACIÓN

MINISTERIO DE ECONOMÍA Y HACIENDA

Buscador de trámites del 060: el Buscador inteligente que integra todos los trámites de la AA.PP de España

GOBIERNO DE ESPAÑA
MINISTERIO DE ECONOMÍA Y HACIENDA

Diseño del Nuevo Portal Semántico de Indicadores e Informes Macroeconómicos del Ministerio de Economía y Empresa

Gobierno de La Rioja

Grafo de conocimiento del Gobierno de La Rioja

LA RIOJA
TURISMO

La Rioja en la Web: Un destino turístico inteligente sobre una Plataforma de IA semántica

GOBIERNO DE ARAGÓN
Departamento de Educación, Universidad, Cultura y Deporte

La plataforma semántica de datos abiertos del Gobierno de Aragón: Aragón Open Data Pool

INSTITUTO GEOGRÁFICO NACIONAL

Mapas y contenidos cartográficos conectados en un gran grafo de conocimiento geográfico

EDUCACIÓN y UNIVERSIDAD

EUSKO JAURLARITZA
GOBIERNO VASCO

Amarauna: la plataforma inteligente de contenidos e innovación educativa de la Consejería de Educación del Gobierno Vasco

SANTILLANA

Set Veintiuno: la plataforma inteligente de Grupo Santillana para educar el talento y las habilidades del siglo XXI

/D/

Didactalia.net, ecosistema educativo de aprendizaje social con capacidades de IA

Junta de Comunidades de Castilla-La Mancha

Plataforma semántica y social para gestionar la formación distribuida del profesorado de Castilla - La Mancha

Telefonica FUNDACIÓN

Plataforma Educativa Profuturo destinada a impulsar su modelo de educación digital inclusiva, que integra contenidos, experiencias educativas y metodologías

Gobierno de La Rioja

Educere: plataforma de recursos educativos del Gobierno de La Rioja

crue Universidades Españolas

HÉRCULES: Sistema Semántico de Gestión de Investigación de las universidades españolas: módulos de Arquitectura Semántica e Infraestructura Ontológica (ASIO) y SGI

Deusto
Universidad de Deusto

Deusto Knowledge Hub, un sistema de interrogación unificado para acceder a toda la información de la universidad

UNIVERSIDAD PONTIFICIA COMILLAS

Comillas Knowledge Hub: buscador unificado del repositorio de docencia e investigación

Comunidad de Madrid
CONSEJERÍA DE EDUCACIÓN E INVESTIGACIÓN

Buscador de Centros Educativos de la Comunidad Autónoma de Madrid

BANCA

BBVA

Buscador BBVA.com, de BBVA Research así como la Plataforma de Reputación del Banco

Triodos Bank

Plataforma de Gestión del Conocimiento para operar en el día a día

EMPRESA

COMUNICACIÓN

CORPORATE EXCELLENCE
CENTRE FOR REPUTATION LEADERSHIP

Gestión inteligente de intangibles para los directivos de comunicación del IBEX 35

RETAIL

adveo

El corazón inteligente del sistema de información de Adveo

SALUD

PSN

Plataforma Inteligente para la gestión de los procesos sanitarios de Prevención Sanitaria Nacional

INNOVACIÓN

COTEC

Plataforma Digital de COTEC: en la web

FORMACIÓN CORPORATIVA

Unikemia
DRIVE FORWARD TRANSFORM

Plataforma E-learning social en un ambiente de IA para la universidad corporativa

gnos

Deusto
Universidad de Deusto
Deusto University
University of Deusto

La Universidad de Deusto, reconocida recientemente como un International Excellence Campus, se fundó en 1886 y cuenta con 6 facultades: Ciencias Sociales y Humanas, Derecho, Ingeniería, Psicología y Educación, Teología y Deusto Business School.

**Deusto – MORElab Research Group:
ICT (Information and Communication Technologies) for Good by realizing Internet of People, where IoP = Hybrid Intelligence + IoT + Human/Machine interaction**

Our main research areas are:

- ☐ Context-aware Mobile Computing for Enhanced User-Environment Interaction
- ☐ Semantic Middleware for Embedded Wirelessly-connected Devices
- ☐ Smart Environments of Augmented Internet-connected Persuasive Objects
- ☐ Ambient Assisted Environments (homes, cities): adaptive accessible services aware of user activities
- ☐ Future Internet: Internet of Services, Internet of Things and Semantic Web
- ☐ ICT applied to Societal Challenges: ageing, open government and sustainability



Past relevant projects:

- **EDI: European Data Incubator**, ICT-14-2017-IA, Pr. 779790 & **REACH: EuRopEAn incubator for trusted and secure data value Chains**, DT-ICT-05-2020 IA: Big Data Innovation Hubs, Pr. 951981. Research areas tackled: Big Data analytics, entrepreneurship, piloting.
- **WeLive: A neW concept of public administration based on citizen co-created mobile urban services**, H2020-INSO-1-2014, 645845. Research areas tackled: co-creation methodology, cobusiness models, exploitation strategies, innovation management.
- **SIMPATICO: SIMplifying the interaction with Public Administration through Information technology for Citizens and cOmpanies**, H2020-EURO-6-2015, 692819. Research areas tackled: User-generated data, crowdsourcing, NLP, e-services, piloting.
- **City4Age: Elderly-friendly City services for active and healthy ageing**, PHC-21- 2015 Risk detection RIA, 689731. Research areas tackled: activity monitoring, Linked Data, reasoning mechanisms.
- **GreenSoul: Eco-aware Persuasive Networked Data Devices for User Engagement in Energy Efficiency**, H2020-EE-2015-2- RIA, Pr. number: 696129. Research areas tackled: Behaviour change, persuasive computing, IoT, middleware
- **AURORAL: Architecture for Unified Regional and Open digital ecosystems for Rural Areas Large scale application**, DT-ICT-09-2020 IA, 48 meses, Pr. 101016854: Research areas tackled: semantic data interoperability, piloting, Smart Villages

FONDO EUROPEO DE DESARROLLO REGIONAL (FEDER)

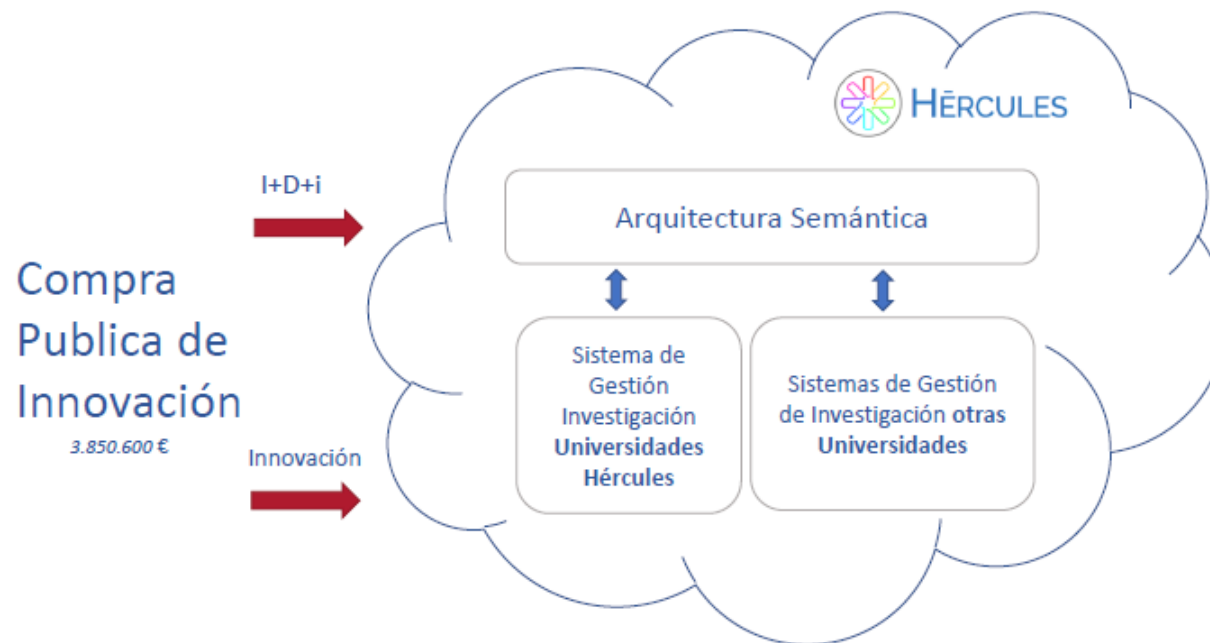
Una manera de hacer Europa

Proyecto Hércules



Hércules. Semántica de Datos de Investigación de Universidades

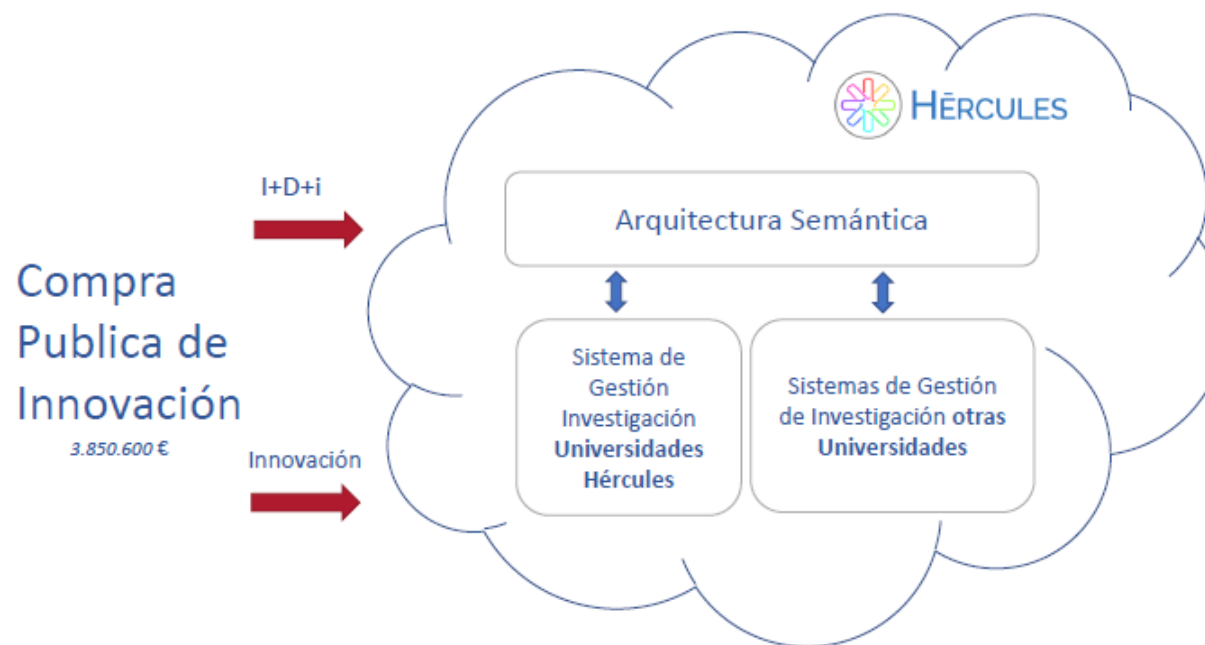
Creación de un Sistema de Gestión de Investigación (SGI) basado en datos abiertos semánticos que ofrezca una visión global de los datos de investigación del Sistema Universitario Español (SUE), para mejorar la gestión, el análisis y las posibles sinergias entre universidades y el gran público (<https://www.um.es/web/hercules/sobre-hercules>)



Hércules. Semántica de Datos de Investigación de Universidades

Estructuración en torno a dos pilares:

- ❑ **Arquitectura Semántica.**
- ❑ Prototipo innovador de un **Sistema de Gestión de la Investigación (SGI).**



Hércules. Semántica de Datos de Investigación de Universidades

El proyecto se divide en **3 subproyectos**:

1. **ASIO**. Arquitectura Semántica de Datos del SUE e Infraestructura Ontológica de la Información del SUE
2. **SGI**. Sistema de Gestión de la Investigación.
3. **EDMA**. Enriquecimiento de Datos a partir de Internet y Desarrollo de Métodos de Análisis Semántico

Hércules. Semántica de Datos de Investigación de Universidades

El proyecto se divide en 3 subproyectos:

1. **ASIO**. Arquitectura Semántica de Datos del SUE e Infraestructura Ontológica de la Información del SUE.
 - Desarrollo una plataforma eficiente para almacenar, gestionar y publicar los datos del SGI, basándose en la Infraestructura Ontológica, con la capacidad de sincronizar instancias instaladas en diferentes Universidades.
 - Creación una red de ontologías que pueda ser usada para describir con fidelidad y alta granularidad los datos del dominio GI (Gestión de la Investigación).
2. SGI. Sistema de Gestión de la Investigación.
3. EDMA. Enriquecimiento de Datos a partir de Internet y Desarrollo de Métodos de Análisis Semántico.

Hércules. Semántica de Datos de Investigación de Universidades

El proyecto se divide en 3 subproyectos:

1. ASIO. Arquitectura Semántica de Datos del SUE e Infraestructura Ontológica de la Información del SUE.
2. **SGI**. Sistema de Gestión de la Investigación.
 - Desarrollo de un prototipo innovador de un sistema de gestión integral de la investigación susceptible de poder ser integrado en cualquier universidad CRUE. Incluirá todas las actividades de gestión relacionadas con la investigación y deberá estar integrado con la arquitectura semántica desarrollada.
3. EDMA. Enriquecimiento de Datos a partir de Internet y Desarrollo de Métodos de Análisis Semántico

Hércules. Semántica de Datos de Investigación de Universidades

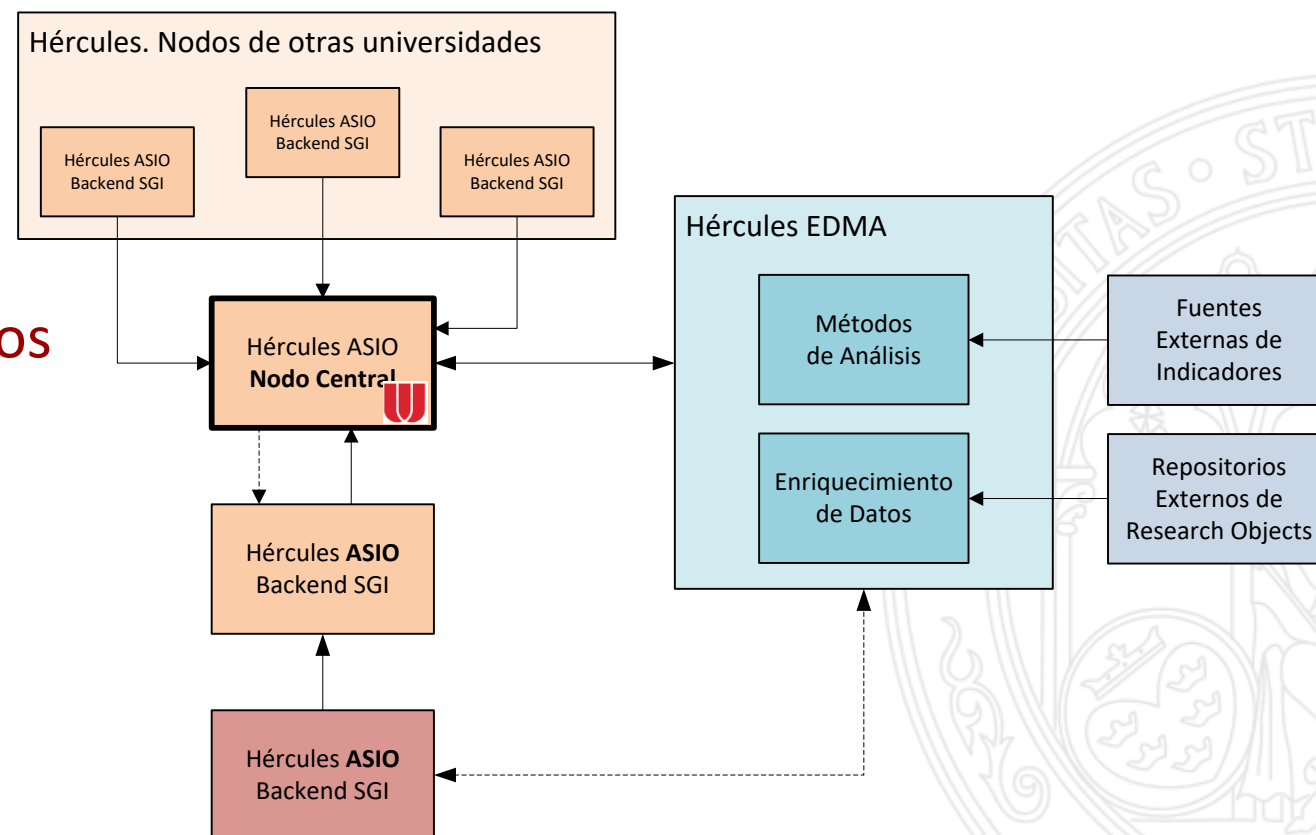
El proyecto se divide en 3 subproyectos:

1. ASIO. Arquitectura Semántica de Datos del SUE e Infraestructura Ontológica de la Información del SUE.
2. SGI. Sistema de Gestión de la Investigación.
3. **EDMA**. Enriquecimiento de Datos a partir de Internet y Desarrollo de Métodos de Análisis Semántico
 - Identificación, extracción, análisis y evaluación de conjuntos de datos relevantes. Completado de datos a partir de la información de fuentes externas de datos.

Hércules. Semántica de Datos de Investigación de Universidades

Comunicación entre los proyectos:

1. SGI **envía datos** a ASIO.
2. ASIO **consolida datos** en un nodo central Unidata.
3. El nodo central Unidata **unifica datos** homogéneos entre nodos.
4. EDMA **explota los datos** del nodo central y enriquece datos desde fuentes de Internet.
5. EDMA se **conecta con SGI** para la gestión de proyectos y CV.



FONDO EUROPEO DE DESARROLLO REGIONAL (FEDER)

Una manera de hacer Europa

Hércules ASIO. Presentación



Hércules ASIO. Arquitectura Semántica e Infraestructura Ontológica

Construcción y explotación de un **grafo de conocimiento** unificado, interrogable, extensible y expresivo **de la investigación universitaria** del Sistema Universitario Español.

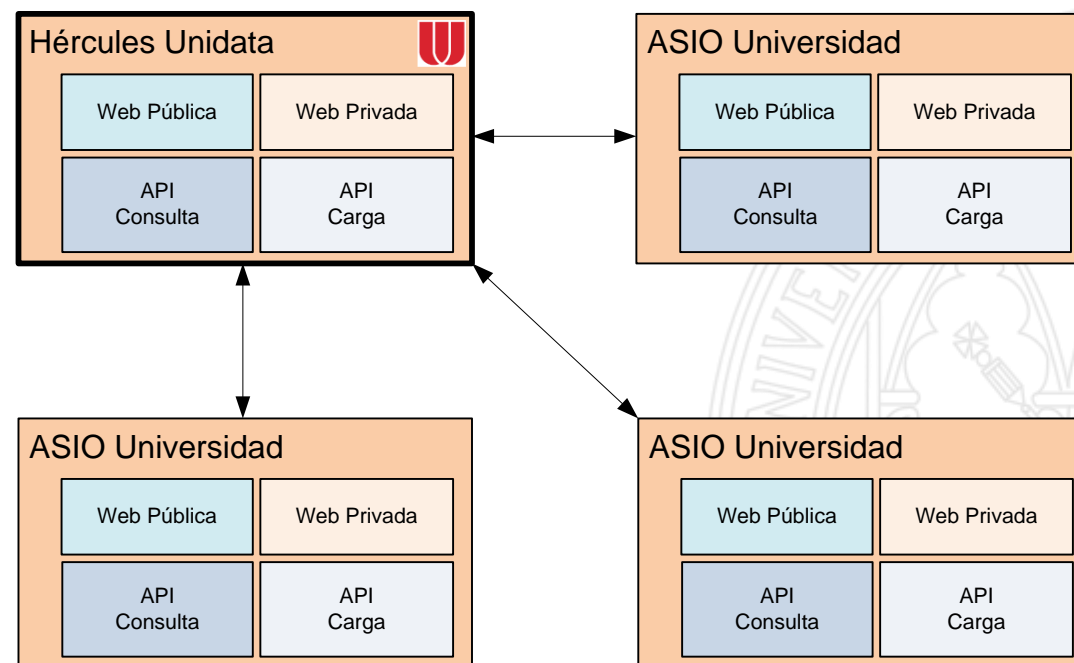
Hércules ASIO **Infraestructura Ontológica**: creación una red de ontologías (ROH) que pueda ser usada para describir con fidelidad y alta granularidad los datos del dominio GI.

Hércules ASIO **Backend SGI Arquitectura Semántica**: es una plataforma eficiente para almacenar, gestionar y publicar los datos del SGI, basándose en la Infraestructura Ontológica, con la capacidad de sincronizar instancias instaladas en diferentes Universidades y permitir explotaciones inteligentes de la información de investigación de la universidad y del SUE.

Hércules ASIO. Arquitectura Semántica e Infraestructura Ontológica

La arquitectura de ASIO se caracteriza por:

1. Cada nodo ASIO tiene 4 módulos: API de **carga**, API de **consulta**, **web pública** y **web privada**.
2. Los nodos ASIO **consolidan datos** en el nodo Unidata.
3. El nodo Unidata **unifica datos** homogéneos entre nodos.
4. Cada nodo ASIO cuenta con un **grafo de conocimiento completo de sus entidades** del dominio GI.
5. El nodo central Unidata cuenta con un **grafo de conocimiento de la globalidad del SUE**, que recibe de cada nodo ASIO.



Hércules ASIO. Arquitectura Semántica e Infraestructura Ontológica

- ❑ ASIO es la **base para para la construcción de un sistema web inteligente** del sistema de investigación de la universidad española.
- ❑ Marco tecnológico:
 - ❑ Programa de **inteligencia artificial interpretada semánticamente**.
 - ❑ Explotación por parte de humanos y máquinas de las posibilidades que ofrece el **enlazado de datos (*linked data*)** en un **grafo de conocimiento (*knowledge graph*)**
- ❑ Infraestructura Ontológica. Definición del **modelo ontológico** del grafo de conocimiento de la investigación universitaria. **Red de Ontologías Hércules (ROH)**
- ❑ Arquitectura Semántica. Desarrollo e implementación de componentes que hacen posible la **carga e interrogación del grafo de conocimiento**.

Hércules ASIO. Datos enlazados y abiertos. Linked Open Data



Hércules ASIO. Datos enlazados y abiertos. Linked Open Data

Los datos abiertos son aquellos que pueden ser utilizados, reutilizados y redistribuidos libremente por cualquier persona, y que se encuentran sujetos, como mucho, al requerimiento de atribución y de compartirse de la misma manera en que aparecen.

Del Open Data Handbook (<http://opendatahandbook.org/guide/es/what-is-open-data/>)

Hércules ASIO. Datos enlazados y abiertos. Linked Open Data

La definición de apertura completa (Open Data Handbook) se resume en:.

- ❑ **Disponibilidad y acceso:** la información debe estar disponible en una forma conveniente y modificable, como un todo y a un costo razonable de reproducción, preferiblemente descargándola de internet.
- ❑ **Reutilización y redistribución:** los términos de uso de los datos deben permitir reutilizarlos y redistribuirlos, e incluso integrarlos con otros conjuntos de datos.
- ❑ **Participación universal:** todos deben poder utilizar, reutilizar y redistribuir la información, sin discriminación en términos de esfuerzo, personas o grupos; ni restricciones “no comerciales” o de uso.

Hércules ASIO. Datos enlazados y abiertos. Linked Open Data

¿Por qué es tan importante ser claros respecto de lo que significa ser “Abierto” y por qué se usa la definición anterior? La respuesta es simple: **interoperabilidad**. Referente a los datos, se trata de la habilidad para interoperar o integrar diferentes conjuntos de datos, permitiendo que componentes distintos trabajen juntos. Así, una parte del material abierto pueda ser mezclado con otro material abierto, lo que resulta fundamental para entender los principales **beneficios prácticos** de la apertura: el incremento dramático de la habilidad de **combinar distintas bases de datos o conjuntos de datos y así desarrollar más y mejores productos y servicios**.

Hércules ASIO. Datos enlazados y abiertos. Linked Open Data

El movimiento de datos abiertos recomienda exponer los datos cuanto antes, sin preocuparse mucho por el formato. Así pues, podrían ser datos abiertos:

- ☐ Un fichero o ficheros **Excel** (es un formato propietario).
- ☐ Un fichero o ficheros en formato **CSV** (tiene problemas con textos con algún formato, como los saltos).
- ☐ Un **API** que proporciona acceso a tablas o vistas desconectadas.
- ☐ **Smart data**. Datos con información acerca de los datos.
- ☐ **Linked Open Data**. Datos abiertos y enlazados (o enlazables) con otros datos.

No todos los datos abiertos son iguales

Hércules ASIO. Datos enlazados y abiertos. Linked Open Data

Es cierto que la **rapidez** es un factor de innovación, pero la **incomodidad** es un freno.

Posibles problemas con los datos:

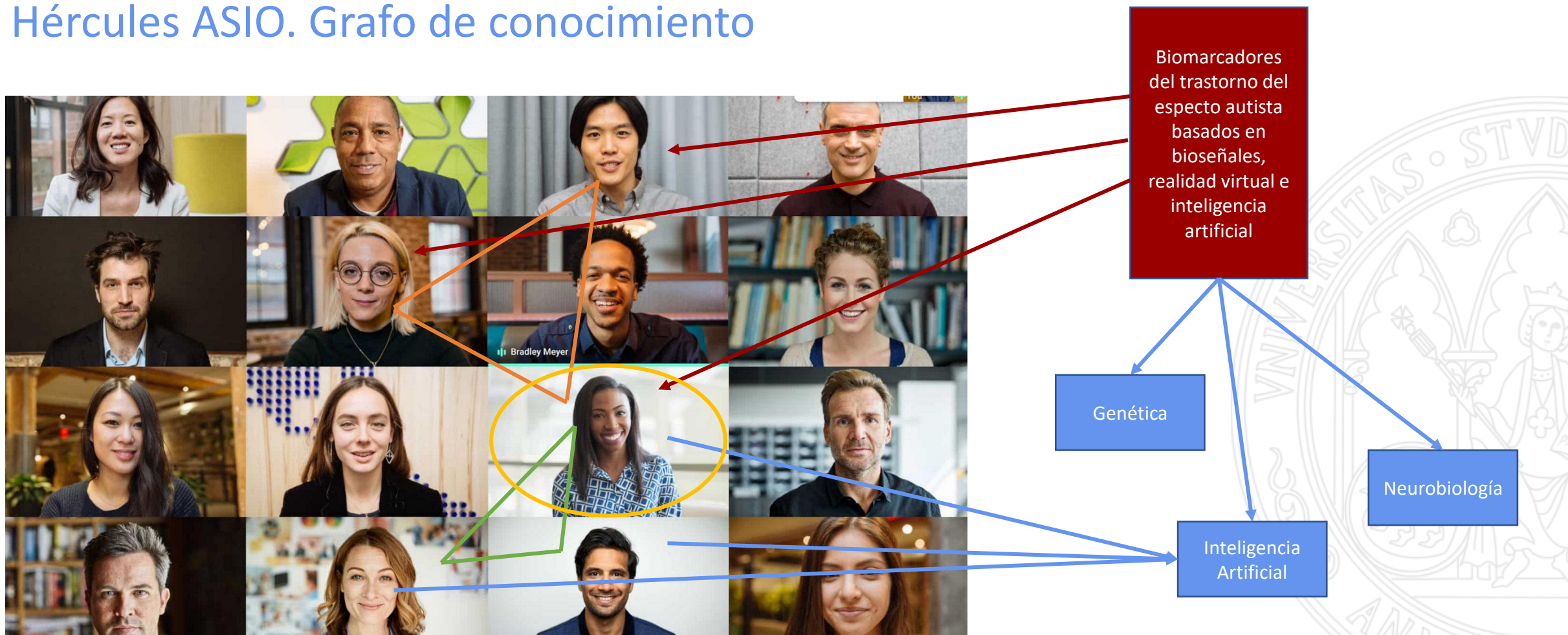
- ☐ **Obtención.** ¿Descarga de ficheros o API?
- ☐ **Modelo.** ¿Cerrado o abierto? ¿Público o privado? ¿Estándar o propietario? ¿Expresivo? ¿Extensible?
- ☐ **Relación.** ¿API sobre vistas o acceso a consultas (como SPARQL o SQL)?
- ☐ **Enlazado.** ¿Relacionable con fuentes externas o cerrado?
- ☐ **Actualización.** ¿Descarga o API? ¿Reutilizador o propietario?
- ☐ **Difusión de los datos.** ¿Enlaces o web? ¿Pasivo o activo?

Linked Open Data es el mejor medio para publicar datos abiertos

Hércules ASIO. Grafo de conocimiento. Knowledge Graph



Hércules ASIO. Grafo de conocimiento



Hércules ASIO. Grafo de conocimiento



Hércules ASIO. Grafo de conocimiento

Las relaciones se pueden expresar en frases simples (predicados de primer orden), por ejemplo:

- ☐ **Laura** es autor/a de la publicación “Biomarcadores del ...”
- ☐ **Genética** es un área de conocimiento de la publicación “Biomarcadores del ...”
- ☐ Neurobiología es un área de conocimiento de la publicación “Biomarcadores del ...”
- ☐ Inteligencia Artificial es un área de conocimiento de la publicación “Biomarcadores del ...”
- ☐ Laura investiga en el área de conocimiento **Inteligencia artificial**
- ☐ Akira es autor/a de la publicación “Biomarcadores del ...”
- ☐ Akira investiga en el área de conocimiento Genética.

...

Y extraer conclusiones/inferencias como: Las personas que investigan en el área de conocimiento de Inteligencia Artificial que han participado en investigaciones en el área de conocimiento de Genética son: Laura, etc.

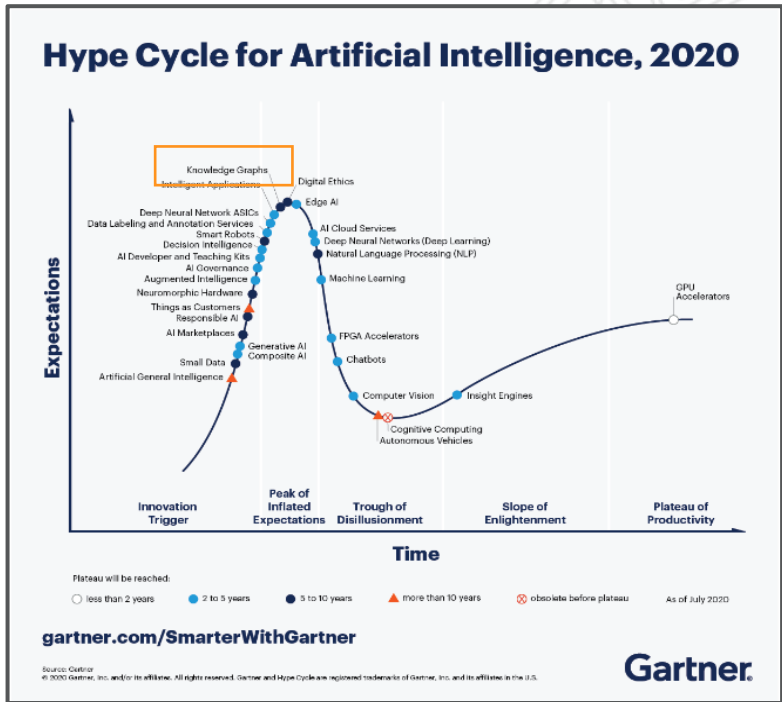
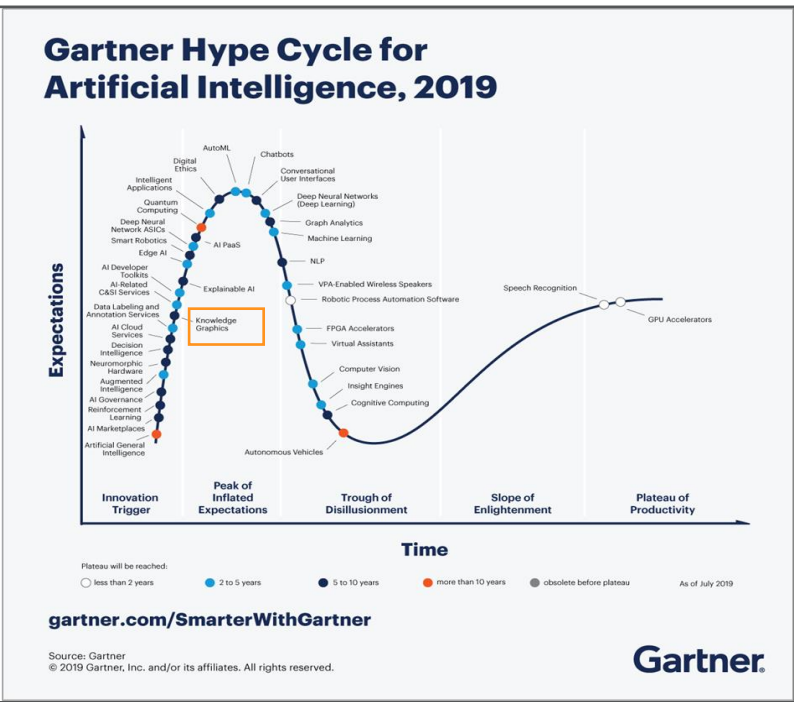
Hércules ASIO. Grafo de conocimiento

Un grafo de conocimiento es una forma de **integrar y representar la información heterogénea y distribuida**, que permite descubrir e investigar cualquier tema de una forma más profunda e intuitiva y disfrutar de una web más semánticamente consciente.

Los grafos de conocimiento permiten **dotar de sentido común** a nuestros sistemas de manera que puedan desarrollar **inteligencia contextual** en un dominio amplio de conocimiento.

Hércules ASIO. Grafo de conocimiento. El futuro de la IA

Gartner sitúa los grafos de conocimiento como una tecnología en crecimiento. Los Grafos de Conocimiento basados en IA se están convirtiendo en parte de las principales tendencias tecnológicas.



FONDO EUROPEO DE DESARROLLO REGIONAL (FEDER)

Una manera de hacer Europa

Hércules ASIO. Infraestructura Ontológica



Hércules ASIO. Infraestructura Ontológica

Definición del modelo ontológico del grafo de conocimiento de la investigación universitaria.

Creación una Red de Ontologías que pueda ser usada para describir con fidelidad y alta granularidad los datos del dominio GI: **Red de Ontologías Hércules (ROH)**.

Hércules ASIO. Infraestructura Ontológica

Especificación de los requerimientos a cumplir por la ROH:

- ❑ **Análisis de escenarios:** mapa del conocimiento nacional, cuadros de mando flexibles, búsqueda de socios, selección de grupos de investigación, etc.
- ❑ **Análisis de funcionalidades del SGI:** proyectos, convocatorias y ayudas, producción científica, CV, contratos y patentes, ética, etc.
- ❑ **Análisis de entidades:** taxonomías, atributos, relaciones y fuentes de datos por cada entidad identificada, como investigador/a, proyecto, fuente de financiación, etc.
- ❑ **Requisitos funcionales y no funcionales de ASIO:** principios de Open Data, principios FAIR, URIs persistentes, multilingüismo, interoperabilidad con otras ontologías, integración con fuentes de información existentes, enlace con grafos de conocimiento externos (linked data), mantenibilidad y licencia Open Source.

Hércules ASIO. Infraestructura Ontológica

Principios seguidos en el diseño de la ROH:

- ❑ **Reutilización**: modelado de conceptos utilizando ontologías y vocabularios existentes y consolidados.
- ❑ **Extensión**: modelado de nuevos conceptos identificados en el análisis o de conceptos propios de la universidad española.
- ❑ **Usabilidad**: documentar los atributos, relaciones y restricciones para facilitar el uso de la ROH por parte de desarrolladores y usuarios.



Hércules ASIO. Infraestructura Ontológica. Desarrollo y mantenibilidad.

El desarrollo de ROH y su mantenibilidad se sustenta en:

- ☐ Principios de diseño: reutilización, extensión y usabilidad.
- ☐ Documentación.
 - ☐ [Documentación descriptiva.](#)
 - ☐ [Documentación para reutilizadores.](#)
 - ☐ [Documentación formal de publicación de ontologías.](#)
- ☐ [Preguntas de competencia](#): un conjunto de consultas extensible que la ontología debe ser capaz de satisfacer, de modo que las posibles modificaciones ontológicas no supongan pérdida de funcionalidad.

Hércules ASIO. Infraestructura Ontológica. Principios FAIR

FAIR (**F**indability, **A**ccessibility, **I**nteroperability and **R**euse). Fair: justo, bueno, adecuado.

[FAIR Guiding Principles for scientific data management and stewardship](#). Scientific Data (2016)

Los Principios FAIR proporcionan **directrices para la publicación de recursos digitales** tales como conjuntos de datos, códigos, flujos de trabajo y objetos de investigación, de manera que sean localizables, accesibles, interoperables y reutilizables (FAIR).

Los principios se refieren a tres tipos de entidades: **datos** (o cualquier objeto digital), **metadatos** (información sobre ese objeto digital) e **infraestructura**.

Linked Open Data es el mejor medio para ser FAIR Data

Hércules ASIO. Infraestructura Ontológica. Principios FAIR.

Findable (encontrable / localizable): los metadatos y los datos deben ser fáciles de encontrar tanto para los seres humanos como para los ordenadores. Los metadatos legibles por máquinas son esenciales para el descubrimiento automático de conjuntos de datos y servicios. Los principios FAIR relativos a *Findable* son:

- ☐ F1. A los (meta)datos se les asigna un identificador global único y persistente.
- ☐ F2. Los datos se describen con metadatos ricos (definidos por R1)
- ☐ F3. Los metadatos incluyen clara y explícitamente el identificador de los datos que describen.
- ☐ F4. Los (meta)datos se registran o indexan en un recurso de búsqueda.

Linked Open Data proporciona un URI, metadatos descriptivos,
identificador y un medio de registro e indexación

Hércules ASIO. Infraestructura Ontológica. Principios FAIR.

Accesible: una vez que el usuario encuentra los datos requeridos, necesita saber cómo se puede acceder a ellos, posiblemente incluyendo autenticación y autorización. Los principios FAIR relativos a *Accesible* son:

- ☐ A1. Los (meta)datos son recuperables mediante su identificador utilizando un protocolo de comunicaciones normalizado.
 - ☐ A1.1 El protocolo es abierto, gratuito y de aplicación universal
 - ☐ A1.2 El protocolo permite un procedimiento de autenticación y autorización, cuando sea necesario
- ☐ A2. Los metadatos son accesibles, incluso cuando los datos ya no están disponibles

Linked Open Data da acceso a los datos mediante el URI, con un protocolo abierto (http) y proporciona persistencia, incluso en borrados

Hércules ASIO. Infraestructura Ontológica. Principios FAIR.

Interoperable: por lo general, los datos deben integrarse con otros datos. Además, los datos deben interoperar con aplicaciones o flujos de trabajo para su análisis, almacenamiento y procesamiento. Los principios FAIR relativos a *Interoperable* son:

- ☐ I1. Los (meta)datos utilizan un lenguaje formal, accesible, compartido y ampliamente aplicable para la representación del conocimiento.
- ☐ I2. Los (meta)datos utilizan vocabularios que siguen los principios de FAIR.
- ☐ I3. Los (meta)datos incluyen referencias cualificadas a otros (meta)datos.

Linked Open Data representa los datos mediante ontologías; reutiliza o propone vocabularios estándar, abiertos y accesibles; y permite enlazar con otros datos

Hércules ASIO. Infraestructura Ontológica. Principios FAIR.

Reuse (Reutilizable): el objetivo final de FAIR es optimizar la reutilización de los datos. Para lograr esto, los metadatos y los datos deben estar bien descritos para que puedan ser replicados y/o combinados. Los principios FAIR relativos a *Reuse* son:

- ☐ R1. Los meta(datos) se describen ampliamente con una pluralidad de atributos precisos y relevantes
 - ☐ R1.1. Los (meta)datos se liberan con una licencia de uso de datos clara y accesible.
 - ☐ R1.2. Los (meta)datos están asociados a una procedencia detallada
- ☐ R1.3. Los (meta)datos cumplen con los estándares de la comunidad relevantes para el dominio.

Linked Open Data en ASIO utiliza vocabularios ricos y descriptivos, que incluyen licencia y procedencia; adecuados para el dominio GI

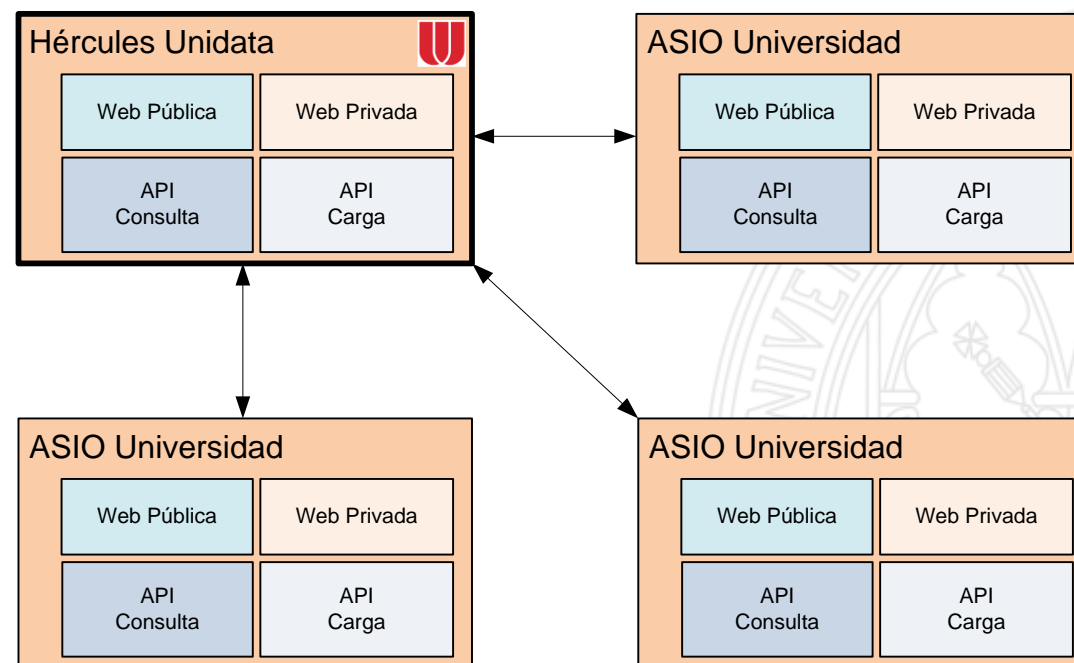
Hércules ASIO. Arquitectura Semántica



Hércules ASIO. Arquitectura Semántica

La arquitectura de ASIO se caracteriza por:

1. Cada nodo ASIO tiene 4 módulos: API de **carga**, API de **consulta**, **web pública** y **web privada**.
2. Los nodos ASIO **consolidan datos** en el nodo Unidata.
3. El nodo Unidata **unifica datos** homogéneos entre nodos.
4. Cada nodo ASIO cuenta con un **grafo de conocimiento completo de sus entidades** del dominio GI.
5. El nodo central Unidata cuenta con un **grafo de conocimiento de la globalidad del SUE**, que recibe de cada nodo ASIO.



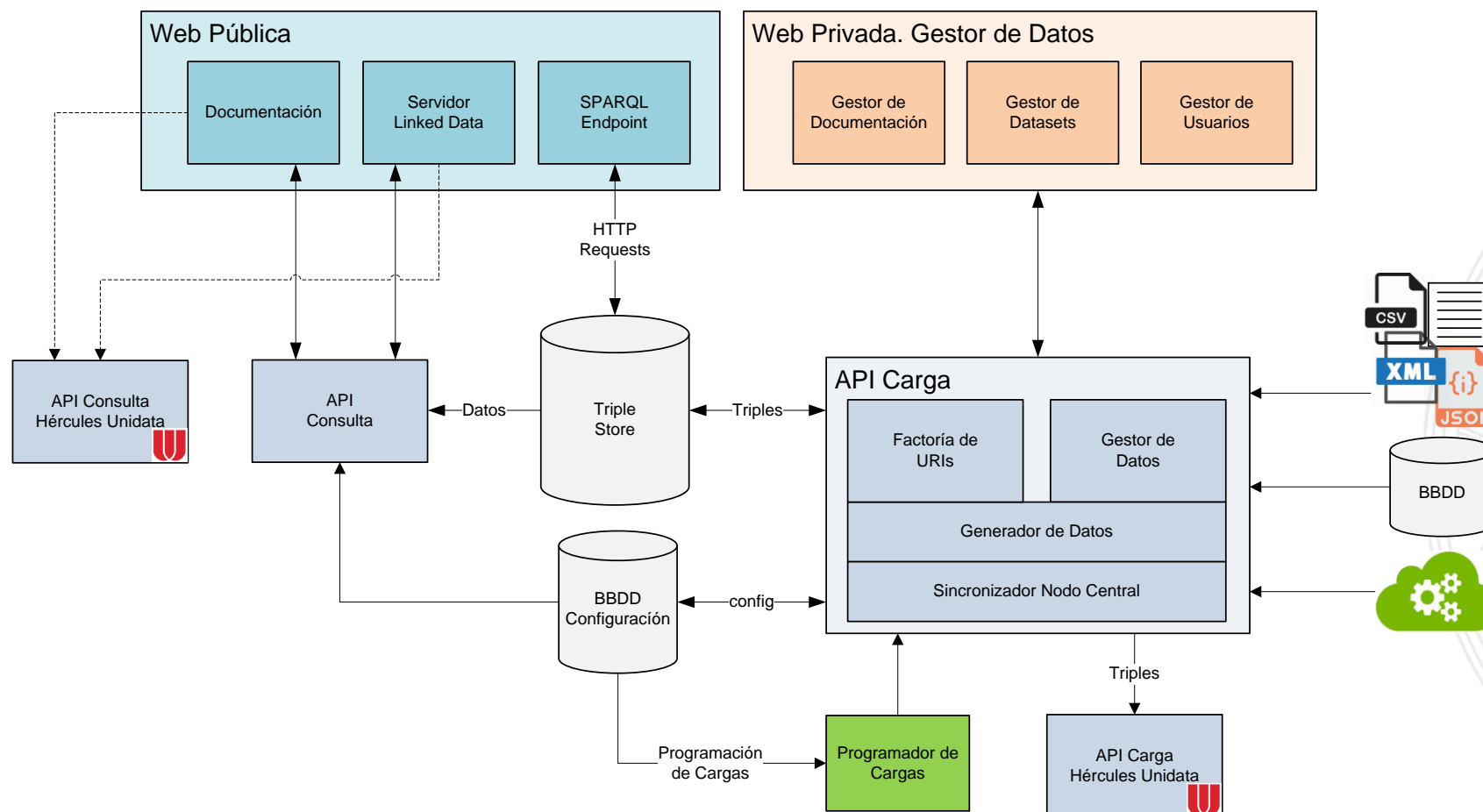
Hércules ASIO. Arquitectura Semántica

El nodo central Unidata cuenta con las siguientes funciones específicas:

- ❑ **Sincronización** con los nodos distribuidos de cada universidad.
- ❑ **Explotación agregada** de datos: unicidad, persistencia, disponibilidad y rendimiento.
- ❑ **Conexión** de datos y **detección de equivalencias** entre universidades.

Sin Unidata, las funciones anteriores tendrían que realizarse mediante consultas a cada uno de los nodos y darían problemas de rendimiento, disponibilidad, consistencia y completitud.

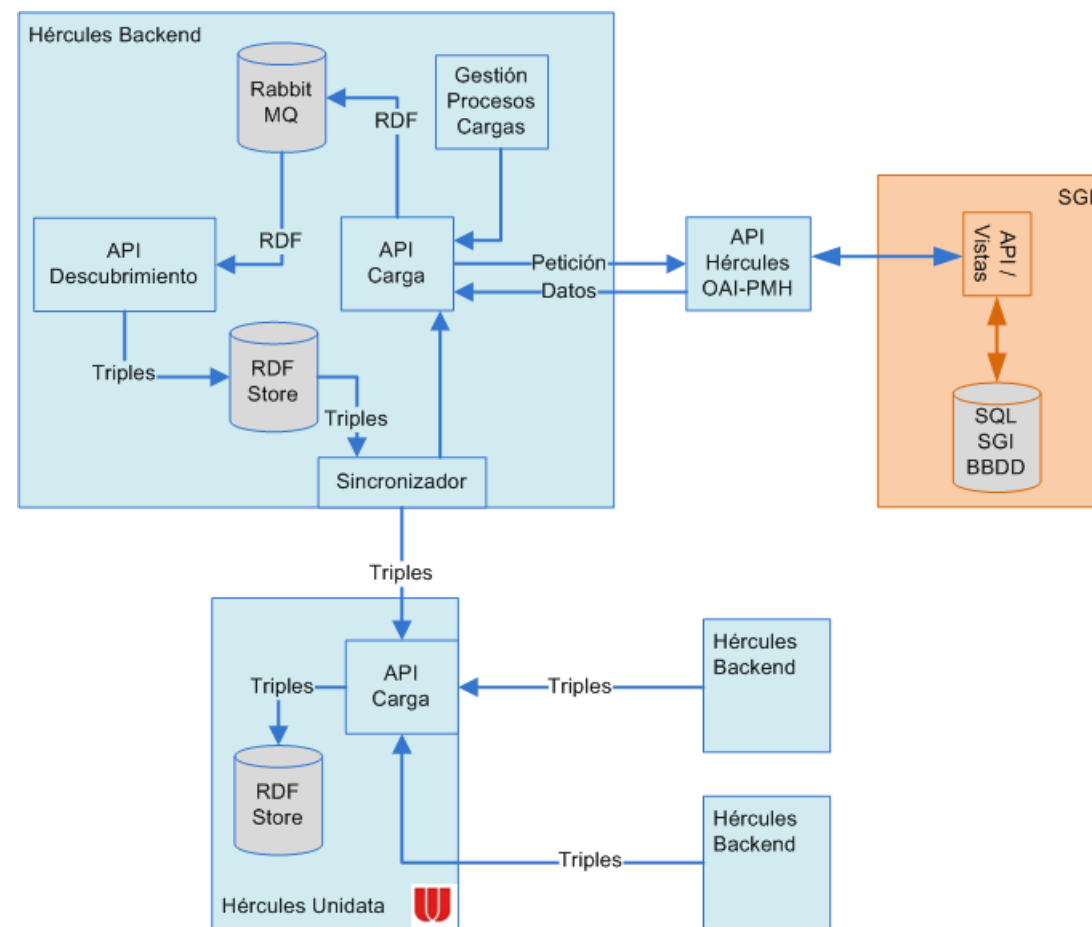
Hércules ASIO. Arquitectura Semántica. Arquitectura de componentes



Hércules ASIO. Arquitectura Semántica. Carga, validación, descubrimiento

El proceso de carga desde el SGI es:

1. Petición de datos nuevos y modificados a un **servicio OAI-PMH** que interroga al SGI.
2. Recepción de los datos en el **API Carga**, en un formato XML definido por un XSD.
3. **Validación** de datos en el API de carga contra la ontología ROH y se **conversión** en RDF.
4. **Descubrimiento** de enlaces, **reconciliación** con los datos existentes y **detección de equivalencias** en el API Descubrimiento.
5. **Publicación** de datos RDF completados en el RDF Store.
6. **Envío** de datos RDF completados al nodo **Unidata**, que cuenta con sus propios procesos de transformación y validación.



Hércules ASIO. Arquitectura Semántica.

Tenemos los siguientes componentes destacados:

- ☐ Carga desde OAI-PMH.
- ☐ Factoría de URIs.
- ☐ Validación.
- ☐ Descubrimiento.
- ☐ Linked Data Server.
- ☐ RDF Store y SPARQL Endpoint

Y los siguientes aspectos a comentar en cuanto a marco tecnológico:

- ☐ Tecnología.
- ☐ Repositorio de Software
- ☐ Despliegues

Hércules ASIO. Arquitectura Semántica. Integración con SGI

Hércules ASIO puede conectarse con cualquier SGI, preferentemente Hércules SGI, mediante:

- ☐ Peticiones a un servicio OAI-PMH ([Open Archive Initiative – Protocol for Metadata Harvesting](#)), cuya fuente de datos sería un SGI.
- ☐ Formato XSD adaptable a diversas fuentes de datos.
- ☐ Conexión con los procesos de ASIO: carga, descubrimiento, validación y publicación.
- ☐ Definición de múltiples fuentes de datos.

Este sistema también permite una carga inicial de datos.

Hércules ASIO. Arquitectura Semántica. URIs. Factoría de URIs

Los datos RDF deben generarse con URIs formalmente correctos, homogéneos y **persistentes**, incluso más allá de la vida del proyecto ASIO. Para ello, contamos con:

- ☐ Documento de **buenas prácticas** de URIs.
- ☐ Especificación de **Esquema de URIs Hércules**, inspirado por la Norma Técnica de Interoperabilidad de Reutilización de recursos de la información (BOE núm. 54, de 4 de marzo de 2013)
- ☐ **API URIs Factory**, que devuelve un URI adecuado para cada tipo de entidad:
 - ☐ `http://graph.um.es/res/researcher/{id-investigador}`
 - ☐ `http://graph.um.es/kos/research-area/{id-categoría}`
- ☐ Documento de **adhesión a las buenas prácticas de URIs**, que se debería aceptar y publicar en el nodo ASIO de cada universidad.

Hércules ASIO. Arquitectura Semántica. Validación

El proceso de validación se efectúa mediante:

- ☐ Uso de *shapes* SHACL ([Shapes Constraint Language. W3C Recommendation](#))
 - ☐ Un lenguaje para validar grafos RDF contra un conjunto de condiciones.
- ☐ Un *conjunto de validaciones*, que se puede relacionar con una fuente de datos.
 - ☐ Podría suceder que ASIO tuviera más de una fuente de datos, de distinta naturaleza o calidad. Por ejemplo, una universidad que no tuviera Hércules SGI y que contase con varias fuentes de datos de varios subsistemas: CV, gestión, etc.
- ☐ Los *shapes* se generan automáticamente *a partir de las restricciones de la ROH* y se pueden modificar manualmente.
 - ☐ El usuario administrador cuenta con herramientas para gestionar los *shapes*.

Hércules ASIO. Arquitectura Semántica. Descubrimiento

Una vez cargados los datos desde el SGI, se ejecutan los siguientes procesos de descubrimiento:

- ☐ Reconciliación.
- ☐ Descubrimiento de enlaces.
- ☐ Detección de equivalencias.

Hércules ASIO. Arquitectura Semántica. Descubrimiento

Una vez cargados los datos desde el SGI, se ejecutan los siguientes procesos de descubrimiento:

- ☐ Reconciliación.
 - ☐ Evita la duplicación de entidades mediante un conjunto de reglas.
 - ☐ Toma decisiones autónomas si la evaluación de las reglas supera un umbral.
 - ☐ Solicita la validación del usuario si la evaluación queda en un rango de duda.
 - ☐ Utiliza datos obtenidos en el descubrimiento de enlaces, desde fuentes externas (ORCID, DOI, etc.) y desde Unidata.
- ☐ Descubrimiento de enlaces.
- ☐ Detección de equivalencias.

Hércules ASIO. Arquitectura Semántica. Descubrimiento

Una vez cargados los datos desde el SGI, se ejecutan los siguientes procesos de descubrimiento:

- ☐ Reconciliación.
- ☒ Descubrimiento de enlaces.
 - ☐ Obtención de identificadores.
 - ☐ Enriquecimiento con enlaces a fuentes externas y/o Unidata.
 - ☐ Información para la reconciliación.
 - ☐ Proceso con ejecución continua que no se ejecuta sólo en el proceso de carga.
- ☐ Detección de equivalencias.

Hércules ASIO. Arquitectura Semántica. Descubrimiento

Una vez cargados los datos desde el SGI, se ejecutan los siguientes procesos de descubrimiento:

- ☐ Reconciliación.
- ☐ Descubrimiento de enlaces.
- ☐ Detección de equivalencias.
 - ☐ Obtención de enlaces a entidades de otros nodos ASIO.
 - ☐ Alineamiento de conceptos
 - ☐ Uso del nodo Unidata.
 - ☐ Información para la reconciliación.

Hércules ASIO. Arquitectura Semántica. Linked Data Server

Linked Data en el RDF St las siguientes

- ☐ Proporcion la presencia
- ☐ Proporcion
- ☐ Cumple l
- ☐ Cuenta c

http://graph.um.es/res/article/bae9443b-9d51-43c8-843b-ef56c664b79c	Towards semantic web-based management of security services	http://purl.org/ron/mirror/bibo#AcademicArticle
http://graph.um.es/res/article/6bcd4117-6a6a-4661-b165-5a7b9858aa50	Towards semantic-aware management of security services in GT4	http://purl.org/roh/mirror/bibo#AcademicArticle
http://graph.um.es/res/article/6e10fef0-ae84-43f9-bfc8-b500c16e6316	Toward a Framework for the Specification of Hybrid Fuzzy Modeling	http://purl.org/roh/mirror/bibo#AcademicArticle

Mostrando página 1 de 3

Anterior 1 2 3

Coautores



se cargados
vicio tiene
e incluyen
máquinas.
nombre.

Hércules ASIO. Arquitectura Semántica. SPARQL Endpoint

Además de Linked Data Server, ASIO cuenta con un punto SPARQL (*endpoint*) que permite a usuarios y administradores consultar los datos del grafo de conocimiento almacenados en el RDF Store.

- ❑ SPARQL: SPARQL Protocol and RDF Query Language.
- ❑ Protocolo de consulta y lenguaje de interrogación para grafos RDF, normalizado por el RDF Data Access Working Group del World Wide Web Consortium (W3C).
- ❑ [SPARQL Versión 1.1 de 2013](#).
- ❑ *Endpoint* de sólo lectura para usuarios públicos, sin acceso a los datos privados.

Reutilización: Conocer ROH + SPARQL Queries + Linked Data Server

Hércules ASIO. Arquitectura Semántica. RDF Store

El RDF Store es el componente de la arquitectura que **almacena el grafo de conocimiento de ASIO en formato RDF**. Dentro de ASIO actúa como un servidor de datos que responde a consultas SPARQL, por lo que también puede ser nombrado como **Servidor RDF o Servidor SPARQL**.

El RDF Store proporciona el servicio de datos y consultas al Linked Data Server y al SPARQL Endpoint.

En el desarrollo de ASIO se ha realizado un *benchmark* para que cada universidad pueda seleccionar la solución que más le convenga. Uno de los requisitos obligatorios definido por el *benchmark* es que debe **cumplir el estándar SPARQL 1.1**, en particular lo referente a actualizaciones de datos.

Hércules ASIO. Arquitectura Semántica. Web Documental

ASIO cuenta con una gestión de web documental que permite publicar páginas webs que informen del nodo ASIO de la universidad, con contenido estático y dinámico, obtenido éste último del API de consulta o del SPARQL Endpoint.


FONDO DE DESARROLLO REGIONAL (FEDER)
Una manera de hacer Europa


HÉRCULES Información sobre Hércules Grafo de Conocimiento Administración

ARQUITECTURA SEMÁNTICA E INFRAEST

ACTUACIONES FINANCIACIÓN FEDER
COMPRA PÚBLICA DE INNOVACIÓN

Hércules ASIO - Arquitectura Semántica Infraestructura Onto

 **HÉRCULES**
Semántica de Datos de Investigación de Universidades

 **crue**
Universidades Españolas

Proyecto Hércules ASIO - Backend SGI

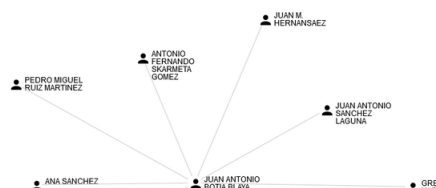
El objetivo del Proyecto ASIO es adquirir un servicio de I+D que consiste en el desarrollo de soluciones innovadoras para la Universidad de Murcia en relación al reto de Arquitectura Semántica e Infraestructura Ontológica. En concreto, se pretende desarrollar e incorporar soluciones que superen las actualmente disponibles en el mercado...[ver más](#)

Grafo de

El proyecto ASIO cuenta con un grafo de 47 triples, generados desde el SGI de la Universidad de Murcia...[ver más](#)

Investigador

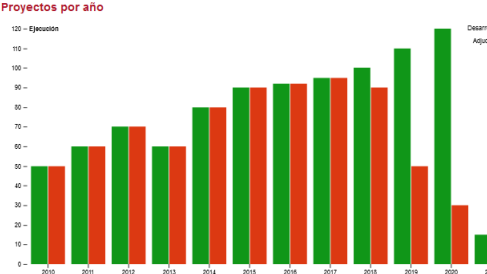
Mapa y grafo de relaciones



FONDO DE DESARROLLO REGIONAL (FEDER)
Una manera de hacer Europa

HÉRCULES Información sobre Hércules Grafo de Conocimiento Administración

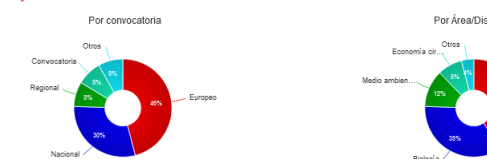
Proyectos por año



Año	Proyectos
2010	50
2011	60
2012	70
2013	60
2014	80
2015	90
2016	90
2017	90
2018	90
2019	100
2020	40
2021	10


Proyectos

Por convocatoria



Convocatoria	Porcentaje
Europeo	40%
Nacional	30%
Regional	20%
Otros	10%

Por Área/Dis



Área/Dis	Porcentaje
Medio ambiente	30%
Biología	20%
Economía	10%
Otros	10%
Humanidades	10%

Datos del Grafo de Conocimiento ASIO-SGI

Los datos del grafo pueden consultarse mediante un SPARQL Endpoint, mientras que sus entidades y relaciones están accesibles como datos enlazados y abiertos (Linked Open Data). Esto convierte los datos de ASIO SGI en un conjunto de datos FAIR.

Consultas SPARQL

SPARQL (SPARQL Protocol and RDF Query Language) es el lenguaje de consulta para RDF y también el protocolo con el que se envían las consultas y se reciben los resultados.


Hércules ASIO dispone de un SPARQL Endpoint público, que permite realizar consultas SPARQL sobre los datos del grafo ASIO-SGI y que se encuentra disponible en la siguiente URL:

<http://155.54.239.204:8890/sparql>

Se indican a continuación algunos ejemplos de consultas:

- Obtener los 100 investigadores con más atributos (más datos).
- Obtener los atributos de un investigador.

Linked Data Server



Linked Data Server es un componente de ASIO desarrollado en Net Core que proporciona el servicio de datos enlazados de Hércules ASIO.

Linked Data Server proporciona acceso a los datos de las entidades del grafo para personas y máquinas, que pueden obtener la representación RDF de la entidad si la petición solicita "application/rdf+xml", y cumple la especificación LDP: Hércules Backend ASIO. Evaluación de cumplimiento Linked Data Platform (LDP)

Se puede comprobar el funcionamiento del servidor mediante los

Hércules ASIO. Arquitectura Semántica. Web de Administración

ASIO proporciona un interfaz de administración, con las siguientes páginas:

- ☐ Administración de repositorios.
- ☐ Gestión de shapes y validación.
- ☐ Factoría de URIs.
- ☐ Obtener tokens de acceso a los APIs.
- ☐ Comprobación del sistema.
- ☐ Administración de páginas de la Web Documental.

Hércules ASIO. Arquitectura Semántica. Tecnología

En el desarrollo de ASIO hemos utilizado las siguientes tecnologías y técnicas de desarrollo:

- ☐ **.Net Core**. Un framework informático administrado, gratuito y de código abierto para los sistemas operativos Windows, Linux y macOS, que sustituye a .NET Framework aportando capacidades multiplataforma. Desarrollado principalmente por Microsoft bajo la Licencia MIT.
- ☐ Arquitectura **SOA (Service Oriented Architecture)**.
 - ☐ Favorece la mantenibilidad y la reutilización.
 - ☐ APIs desarrollados siguiendo el principio “Eat Your Own Dog Food”. ASIO no llama a librerías, sino a los APIs desarrollados e instalados.
 - ☐ Ofrece un modo sencillo de reutilización de componentes de ASIO.
- ☐ Los componentes de ASIO también se pueden reutilizar como librerías en otros desarrollos, pero no es lo que recomendaríamos.

Hércules ASIO. Arquitectura Semántica. Repositorio de Software





El sistema ASIO, en proceso de finalización, está disponible en un repositorio de GitHub, con licencia de software libre:

- ❑ <https://github.com/HerculesCRUE/GnossDeustoBackend>
- ❑ GitHub Actions. Compilación, test unitarios, CodeCov, SonarCloud

🔗 Sobre API CARGA



🔄 Build and test API_CARGA passing  35%

 bugs 6  security C  lines of code 4.7k  duplicated lines 4.4%

Accesible en el entorno de desarrollo en esta dirección a través de swagger: <http://herc-as-front-desata.atica.um.es/carga/swagger/index.html>.

La documentación de la librería está disponible en: http://herc-as-front-desata.atica.um.es/api-carga/library/api/API_CARGA.Controllers.html

[Sin título]



Hércules ASIO. Arquitectura Semántica. Despliegue

El proyecto ASIO se puede desplegar de dos formas:

- ☐ Como un conjunto de Web Services.
- ☐ Mediante imágenes Docker.

La infraestructura básica necesita los siguientes componentes lógicos:

- ☐ Servidor(es) web con Apache.
- ☐ Servidor para tareas de backoffice.
- ☐ RDF Store compatible con SPARQL 1.1
- ☐ Servicio de BBDD relacional (p.e. PostgreSQL) para almacenamiento de opciones de configuración.
- ☐ Servicio de RabbitMQ para la gestión de colas de proceso.

FONDO EUROPEO DE DESARROLLO REGIONAL (FEDER)

Una manera de hacer Europa

GRACIAS

