

Catálogo de soluciones tecnológicas para espacios de datos (Tecnologías, Plataformas y Soluciones)

Impulsando la economía del dato española
mediante la compartición y explotación de datos
de forma confiable, soberana y segura

Fecha: 22/04/2024

Versión 01

En el siguiente documento se muestra un listado de soluciones tecnológicas (Plataformas y Herramientas) útiles para realizar espacios de datos.

***Este es un documento colaborativo que incluye información suministrada por diferentes entidades que han enviado su documentación.**

*Si desean contactar con alguna de las empresas del listado, envíen la petición a la dirección responsable.tecnico@gaia-x.spain.com y les enviaremos la información de contacto.

1. TECNOLOGÍAS, PLATAFORMAS Y SOLUCIONES

ACURATIO

Enlace: <https://www.acuratio.com/>

Sector: *Cualquier sector.*

Plataforma que habilita la compartición de datos para hacer analítica y machine learning preservando la privacidad de los datos. Plataforma premiada por la agencia de protección de datos española. En producción en las entidades financieras más importantes de Europa.

DEDOMENA - Dataverso

Enlace: <https://dedomena.ai/>

Sector: *Cualquier sector; actualmente la mayoría de los datos añadidos provienen de sectores como Banca, Seguros, Telecomunicaciones y Salud.*

El Dataverso es un espacio de datos multisectorial donde empresas, profesionales e individuos pueden encontrar los datos seguros y de calidad que necesitan para innovar, investigar, analizar o construir, mejorar y probar aplicaciones de datos, machine learning e inteligencia artificial. Navega por varias industrias y regiones pudiendo adquirir datos anónimos o sintéticos para diferentes casos de uso, aplicaciones, comportamientos, o fechas específicas.

DIH DATALIFE – TECNOLOGÍA MULTISECTORIAL DATALIFE

Enlace: <https://www.dihdatalife.com/>

Sector: *Cualquier sector.*

Pulse en el siguiente enlace para más información: [DEMOSTRADOR_MULTISECTORIAL_DATALIFE.pdf](#)

Uno de los objetivos que se persigue con la creación del demostrador y que llevará por nombre **demostrador multisectorial DATALife**, es disponibilizar el acceso, los medios de compartición segura y soberana de los datos, además de una infraestructura cloud dotada de modernos y potentes medios de cómputo, tanto a nivel de hardware como de software de toda la cadena de valor, desde un punto de vista holístico que incluyan datos relacionados con la salud humana, la salud animal, el medioambiente (calidad del agua, biodiversidad, climatología, etc.), los biorecursos (mar-pesca, agricultura, ganadería, forestal), los residuos como recursos a valorizar, alimentación, laboratorios control de calidad/analíticos, cultivos y



Asociación Gaia-X España. C.F.: G72514771



652528934



C/ Río Tajo, 2 Talavera de la Reina (Toledo)



gerencia@gaia-x-spain.com

gaiax.es



su aprovechamiento, etc), salud (datos fisiológicos, genéticos, hábitos saludables, salud mental...) y cuidados (parámetros ambientales, hogar, residencias, centros de día...). Además, se vinculará a los recursos de innovación (servicios, infraestructuras y know-how).

El **demostrador multisectorial DATAlife** se articulará de forma modular para poder adecuarse a los sectores más vinculados con los conceptos de OneHealth y bioeconomía. La integración de los principios de OneHealth y la bioeconomía en el desarrollo de este demostrador es esencial para abordar los desafíos actuales en términos de salud, sostenibilidad y desarrollo económico.

El **demostrador multisectorial DATAlife** estará basado en los componentes en desarrollo por el proyecto Eclipse Dataspace Components (EDC) y se suministrará como un servicio empaquetado que proporcionará un entorno seguro y confiable para compartir datos entre participantes autorizados. Estos participantes podrán compartir datos de manera segura y controlar su acceso, garantizando la privacidad y el cumplimiento de regulaciones como el GDPR. El sistema será extensible, lo que permitirá agregar nuevas funcionalidades mediante la implementación de extensiones. Además, cumplirá con los requisitos fundamentales del Anexo VI-Requisitos Técnicos de los Espacios de Datos del BOE número 16, publicado el 18 de enero de 2024.

ELLIOT CLOUD

Enlace: <https://elliottcloud.com/>

Sector: *Multisectorial*.

Elliot Cloud es una plataforma de software low-code IoT orientada a conectar dispositivos en diferentes industrias y desarrollado bajo una plataforma como modelo de servicio. La solución es capaz de cubrir **toda la cadena de valor** de una solución IoT: recolección de datos, gestión de dispositivos, almacenamiento seguro de datos, visualización de datos, definición de flujo de negocios y la integración de ERP, CRM o sistemas de producción ya existentes.

Elliot Cloud facilita el despliegue de dispositivos mediante una interfaz de fácil uso y un modelo drag & drop para generar flujo de datos basados en el dispositivo y la interconexión del sistema. Por lo tanto, la implementación y control de dispositivos son escalables y transparentes para usuarios no técnicos.

Nos gusta presentar a Elliot Cloud como un software de configuración para proyectos IoT con tres pasos: conectar, definir y controlar. Sus principales características son:

- Elliot Cloud es un middleware IOT Open Source con alta escalabilidad y muy rápida integración y con capacidad de conectarse con cualquier sensor, dispositivo, maquinaria o fuente datos existente para la extracción de sus datos.
- Permite gestionar datos en tiempo real con alta variabilidad en el tiempo y una analítica avanzada mediante tecnologías como Big Data e IA.
- Eficiente los procesos a través de analítica avanzada de los datos y reglas de negocio a modelos predictivos nativos para la industria y la empresa conectada.
- Elliot está pensada como una solución abierta, que no limita las barreras de la evolución del entorno tecnológico.



Asociación Gaia-X España. CIF: G72514771



652528934



C/ Río Tajo, 2 Talavera de la Reina (Toledo)



gerencia@gaiax-spain.com

gaiax.es



- Nuestras soluciones están pensadas para eficientar los procesos conectando de forma nativa con todos los sistemas ya existentes.

GRADIANT - AnonimiX

Enlace: <https://www.gradient.org/>

Sector: *Cualquier sector, pero es especialmente útil para los sectores de Salud y de Economía Social y Cuidados.*

La anonimización es un proceso que permite transformar datos personales de forma que éstos no puedan ser vinculados con los individuos a los que hacen referencia. Ayuda a las organizaciones a extraer valor de los datos, mejorando los procesos de toma de decisión y garantizando el cumplimiento normativo (GDPR). Sin embargo, la mayor parte de las soluciones de anonimización actuales presentan bastantes limitaciones, ya que son herramientas complicadas de utilizar y no suelen ofrecer soporte a grandes cantidades de datos.

AnonymiX es una herramienta de anonimización desarrollada por Gradient que ofrece una solución innovadora a estos problemas. La herramienta, compatible con los espacios de datos, permite realizar diversas operaciones de anonimización sobre diferentes tipos de datos estructurados, incluyendo datos procedentes de GPS. La herramienta calcula de forma automática todas las formas posibles de anonimizar los datos, garantizando su privacidad y utilidad, simplificando el proceso al máximo para los usuarios. Dispone de diferentes métricas, que permiten obtener una estimación del riesgo de re-identificación y de la distorsión sufrida por los datos tras su anonimización. Adicionalmente, la herramienta dispone de soporte Big Data y ofrece anonimización en streaming. AnonymiX ofrece diferentes opciones de despliegue, ya que puede usarse como una herramienta on-premise o en la nube, a través de una interfaz web o una API REST.

El nivel de madurez tecnológico (TRL) actual de AnonimiX es 7.

GRADIANT – Clever

Enlace: <https://www.gradient.org/portfolio/clever/>

Sector: *Cualquier sector, pero es especialmente útil para el sector de Industria.*

Clever es una plataforma desarrollada por Gradient que integra análisis de producción, negocios y casos de uso de personas a través de varios módulos. Incorpora los conceptos de Industrial IoT (IIoT) y Digital Twin, lo que le permite procesar grandes cantidades de datos en tiempo real extraídos de diversas fuentes de datos y aplicar algoritmos avanzados y específicos de ML/AI a los mismos. El objetivo final de esta plataforma es mejorar la toma de decisiones y la eficiencia operativa en las empresas donde se implementa.

El nivel de madurez tecnológico (TRL) actual de Clever es 7.



GRADIANT – Plataformas IoT

Enlace: <https://www.gradiant.org/>

Sector: *Cualquier sector.*

Los siguientes plataformas y módulos se pueden componer para crear un producto empaquetado que se puede licenciar.

Gateway industrial para la ingesta de datos de planta. TRL 6/7.

- Un dispositivo que se instala en campo y lee datos de diversas fuentes, los cuales envía a la nube en un formato conocido utilizando el protocolo MQTT. Este dispositivo es apto para Modbus, protocolos de comunicación industriales como Profinet, señales digitales conectadas al dispositivo y, en general, se puede adaptar a cualquier protocolo privativo sobre TCP. Utiliza un dispositivo comercial llamado Revolution PI.

Módulo de software de ingesta de datos usando protocolos a medida. TRL 5-7.

- Proporciona servicios de lectura periódica de datos de diversos protocolos de comunicación, que se pueden adaptar a las necesidades de las fuentes de datos disponibles.
- Se puede usar en el edge o en el cloud, según la conectividad disponible a las fuentes de datos o la aplicación que se busque.
- Se puede configurar para publicar las muestras en un middleware de tipo pub/sub como MQTT o RabbitMQ.
Si el protocolo o formato de datos es nuevo sería necesaria hacer una adaptación para poder leerlo.

Clever Monitor. TRL 6-7.

- Clever es una solución completa, pero modular, con UI, control de acceso y una colección de módulos de funcionalidad. Clever Monitor es uno de estos módulos, en los que se pueden definir de forma jerárquica las entidades que generan datos y configurar lectores OPC-UA para recoger datos continuamente de uno o más servidores OPC-UA. Estos datos se pueden almacenar en base de datos o publicar en RabbitMQ.

APIs y servicios ETL. TRL 5/6.

- Servicios que realizan tareas de ETL, leyendo datos de una base de datos SQL o bus de datos (RabbitMQ). Estos datos los transforman y los guardan transformados en base de datos o pueden publicarlos en RabbitMQ.
- Se puede adaptar la parte de salida de datos para usar otro tipo de tecnologías de streaming, como Kafka.



LoRa data acquisition y streaming. TRL 5-7.

- Módulos software para la ingesta de datos de dispositivos LoRa como estaciones meteorológicas, medidores de humedad de la tierra, sensores de luminosidad, calidad del aire, etc.
- Los datos se tendrían en MQTT, disponibles para usar en streaming o combinar con servicios ETL específicos que preparen los datos para su uso.

GRADIANT - XDataHub

Enlace: <https://www.gradient.org/>

Sector: *Cualquier sector.*

XDataHub es un servicio empaquetado desarrollado por Gradient. Proporciona un entorno seguro y confiable para compartir datos entre participantes autorizados. Los usuarios pueden compartir datos de manera segura y controlar su acceso, garantizando la privacidad y el cumplimiento de regulaciones como el GDPR. El sistema es extensible, lo que permite agregar nuevas funcionalidades mediante la implementación de extensiones.

Como base para el desarrollo, XDataHub utiliza los componentes del proyecto Eclipse Dataspace Components (EDC). Las opciones de implementación incluyen posibilidades locales o en la nube, siendo que la implementación en la nube requiere un componente intermedio, como un proxy inverso, para la comunicación segura.

El sistema cumple con los requisitos fundamentales del Anexo VI-Requisitos Técnicos de los Espacios de Datos del BOE número 16, publicado el 18 de enero de 2024. Sin embargo, actualmente, no proporciona medios para monitorear la interoperabilidad dentro del espacio de datos en tiempo de ejecución (requisito 5).

El nivel de madurez tecnológico (TRL) actual de XDataHub es 3/4.

GRAPHENUS

Enlace: <https://graphenus.io/>

Sector: *Multisectorial.*

Graphenus es una **plataforma SaaS** que permite gestionar los datos, analítica e IA en una única plataforma. Está formada por las mejores herramientas de captación, **almacenamiento, procesamiento, explotación** y consulta de **grandes volúmenes** de datos; **gobernando y securizando el end-to-end** de este ciclo. Además, permite **ingestar** datos desde **diferentes orígenes**, almacenarlos de una forma fiable y tolerante a fallos, realizar analíticas complejas sobre estos (batch, streaming, grafos, ...), asegurar la persistencia de un modelo de datos, utilizar diferentes técnicas de Machine Learning (ML) y representarlos de forma gráfica, para su posterior consulta y explotación.



Se destaca Graphenus para su aplicabilidad en salud por su capacidad de virtualización del dato y la aplicación de tres características distintivas de IA: Aprendizaje Federado, IA Ontológica e IA Generativa, que en conjunto facilitan un importante ahorro de costes, tiempos de programación reducidos y herramientas de optimización de última generación.

En línea con garantizar la gobernanza de datos y la **interoperabilidad** de primer nivel, Graphenus se ha comprometido a **adherirse** al protocolo 'Graphenus Gaia X Ready', enfatizando su dedicación a las mejores prácticas en la gestión y operación de datos.

Graphenus ha sido diseñada y construida en base a una serie de principios que consideramos clave:

- **Multipropósito:** Graphenus dispone de multitud de funcionalidad out-of-the-box, facilitando la implementación de distintos casos de uso, incluyendo analytics y ML.
- **Truly Hybrid:** Puede ser desplegado en infraestructura On Premise o Cloud.
- **Open source:** Graphenus integra y securiza las herramientas open source de tratamiento y analítica de datos más extendidas en sus versiones más recientes.
- **Escalable:** Graphenus está totalmente contenerizado, lo que le permite adaptarse a la demanda prevista y garantiza un rendimiento y una capacidad óptimos, permitiendo escalar los servicios de acuerdo a la necesidad de cada momento.
- **Flexible y modular:** Graphenus permite agregar/eliminar componentes mediante mecanismos plug&play, adaptándose totalmente a cada caso de uso.
- **Seguridad:** Graphenus sigue un principio de confianza cero, en el que todos los permisos deben ser otorgados de forma explícita y se mantiene una auditoría continua de las operaciones realizadas sobre la plataforma.
- **Gobierno:** Graphenus incorpora las herramientas necesarias para garantizar un adecuado gobierno y linaje del dato.
- **Capacidades de ML:** Integración de las principales librerías de Machine Learning.
- **Monitorización unificada:** Recopilación de logs, unificación y monitorización de todos los servicios

ITELLIGENT

Espacio de datos para el vehículo conectado.

Enlace: <https://itelligent.es/>

Sector: *Automoción.*

Desarrollo de un espacio de datos basado en la arquitectura de IDSA. El objetivo de este espacio de datos es que todos los miembros del consorcio con necesidad de compartir información para la ejecución de sus casos de uso, se integren en este espacio de datos para en primer lugar poder obtener la información que necesiten y en un momento posterior poder publicar en él sus modelos como servicio.



ITI - Wisephere

Enlace: <https://www.youtube.com/watch?v=SkLKNVvj3KQ>

Sector: *Cualquier sector.*

Plataforma que convierte los datos de las empresas en información estratégica. Alineado con la Estrategia Europea de Datos y algunas de las principales iniciativas en este campo (GAIA-X, BDVA, IDSA, EUH4D, DSBA, etc.).

MOVEN

Enlace: <https://moven.pro/es/>

Sector: *Movilidad.*

MOVEN ofrece una plataforma que actúa como marketplace de casos de uso de movilidad. Esta plataforma aplica Inteligencia Artificial (IA) para analizar con precisión datos procedentes de diversas fuentes, tanto públicas como privadas al integrarse en un espacio de datos. Nuestra plataforma actúa como interfaz para visualizar los datos y los casos de uso y proporciona a particulares, empresas y ciudades la información necesaria para tomar decisiones informadas en materia de movilidad urbana y planificación del transporte. Nuestro objetivo es hacer que los datos críticos de la movilidad sean accesibles a todo el mundo dando acceso a información de valor procedente de diferentes fuentes de datos gracias a los espacios de datos.

TEKNIKER – Tekniker IDS Connector

Enlace: <https://www.tekniker.es/es>

Sector: *Cualquier sector.*

El Tekniker IDS Connector, siguiendo el Dataspace Protocol ([Dataspace Protocol 2024-1 | Dataspace Protocol | IDS Knowledge Base \(internationaldataspaces.org\)](#)) recientemente lanzado por la International Data Spaces Association ([Home - International Data Spaces](#)), permite a las empresas establecer un único punto de entrada a sus múltiples fuentes de datos, al mismo tiempo que garantiza los 3 pilares fundamentales de un espacio de datos: la interoperabilidad de los datos compartidos en el espacio de datos, la confianza entre las partes involucradas en el intercambio de datos y la soberanía de los datos a lo largo de todo su ciclo de vida. El Tekniker IDS Connector, por tanto, se convierte en un componente técnico core que cualquier empresa debe desplegar para poder participar en un espacio de datos y a partir del cual se pueden construir herramientas que ofrezcan valor añadido.



UDL (Universidad de Lleida)

Enlace: <https://dataspace.angliru.udl.cat/>

Sector: *Agricultura.*

Una plataforma de investigación abierta que sigue los principios FAIR de localizabilidad, accesibilidad, interoperabilidad y reutilización, garantizando al mismo tiempo la soberanía de los datos.

UPC - TECNOLOGÍA UPCxels

Enlace: <https://www.upc.edu/>

Sector: *Cualquier sector con funcionalidad específica para Agroalimentario (UPCxels4Agro), Movilidad (UPCxels4Mobility) y Salud (UPCxels4Health)*

El proyecto demostrador de espacio de datos **UPCxels** se plantea como un ecosistema abierto y multi-sectorial, permitiendo la creación de espacios de datos distintos y de diferentes ámbitos. En el proyecto se pondrá énfasis en el caso de uso Agroalimentario, creando el espacio **UPCxels4Agro** y en él, se enriquecerán datos climáticos, agronómicos de cultivos al aire libre y del resto de la cadena agroalimentaria de diferentes zonas de España con datos de consumo eléctrico, movilidad o zonas de descarga de mercancías, entre otros. Todos los datos se organizarán de acuerdo a su ubicación geográfica. Se plantearán casos de uso en la intersección de diferentes sectores, p. ej. optimización de rutas de distribución de productos frescos, ubicación de plantas de agrivoltaica o predicción de indicadores de salud en función de datos geográficos y climáticos.

Como demostrador abierto y generalizable, UPCxels incluirá también espacios de datos en ámbitos como la movilidad **UPCxels4Mobility**, en la Salud, **UPCxels4Health**, creando la infraestructura especializada, y permitiendo crear servicios también especializados de forma fácil y ágil. Se promoverá la interconectividad y cohesión entre espacios de datos dentro del ecosistema UPCxels.

El espacio de datos UPCxels contará con infraestructura tecnológica propia, e implementará los building blocks propuestos por el DSSC mediante un sistema de gobernanza avanzado propio. Esto permitirá almacenar o acceder remotamente a datos propietarios de empresas, asegurando un marco contractual (mediante smart contracts) seguro para la cesión/compartición de datos garantizando el cumplimiento de las reglas de uso y explotación establecidas, así como un marco tecnológico para la creación y gestión de tecnologías de análisis y enriquecimiento de los datos del espacio. UPCxels se basará en los conceptos generados por Gaia-X y adaptará los últimos estándares e implementaciones abiertas a nivel europeo en privacidad, confianza y gobernanza del dato.



UPM - Incubadora de Espacio de Datos INESData

Enlace: <https://inesdata-project.eu/>

Sector: *Lenguaje, Legal / Contratación pública, Movilidad, Medios.*

INESData (<https://inesdata-project.eu/>) es una Incubadora de Espacios de Datos para fomentar la adopción y acelerar el desarrollo de un ecosistema de Espacios de Datos en España. Se centra en simplificar la adopción de tecnología y acelerar el despliegue industrial del ecosistema nacional de espacios de datos contribuyendo con cuatro espacios de datos (idioma, movilidad, licitación pública y legal, y medios) para demostrar los beneficios de los espacios de datos y la aplicabilidad de la tecnología relacionada.

