



Oferta Tecnológica de Gradiant – XDataHub y AnonimiX

3 de Abril de 2024

Helena Fernández, hfernandez@gradiant.org

(+34) 986 120 430 | gradiant@gradiant.org | www.gradiant.org



Contenido

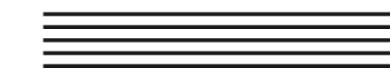
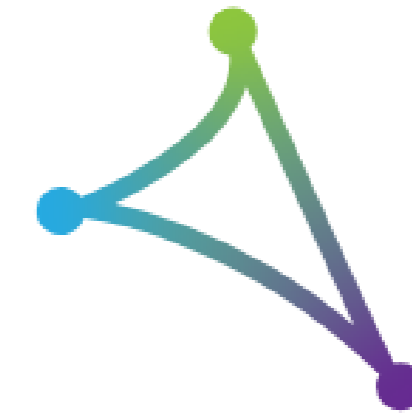


- Proyecto EDC
- Oferta tecnológica de Gradient
 - XDataHub
 - AnonimiX
- XDataHub y AnonimiX: posible despliegue

Proyecto Eclipse Dataspace Components (EDC)



- Código abierto.
- Basado en Java.
- Arquitectura extensible, escalable y personalizable.
- Versión actual: 0.6.0.
- Alineado con IDSA y Gaia-X.
- Comunidad en crecimiento y adopción.



EDC – Principales componentes



Componente	Descripción
Connector	<ul style="list-style-type: none">• Se utiliza cada vez que cambia la entidad controladora (legal) de los datos.• Proporciona una forma genérica de expresar, negociar y documentar las reglas según las cuales se comparten los datos, y también con quién.
Identity Hub	<ul style="list-style-type: none">• Almacena y entrega credenciales de manera segura.• Gestionado por cada uno de los Participantes del espacio de datos.
Federated Catalog	<ul style="list-style-type: none">• Nodo de Catálogo Federado: Publica activos, permitiendo su descubrimiento. Un activo puede estar disponible para un subconjunto de participantes mediante una política de acceso. Se puede imponer requisitos de uso con una política de uso.• Explorador de Catálogo Federado: Un sistema que descubre activos publicados por otros participantes en un espacio de datos.
Registration Service	<ul style="list-style-type: none">• Componente clave de la Autoridad del Espacio de Datos (Data Space Authority).• Permite que el Participante registre su DID y emite una Credencial Verificable (VC) de membresía al Participante. Además, permite la verificación de credenciales.
Data Dashboard	<ul style="list-style-type: none">• Interfaz de gestión.

Oferta tecnológica: XDataHub



- **Servicio empaquetado** que proporciona un entorno seguro y confiable para compartir datos entre participantes autorizados. Garantiza la privacidad y el cumplimiento de regulaciones como el GDPR.
- **Basado en EDC, siendo extensible**, lo que permite agregar nuevas funcionalidades mediante la implementación de extensiones.
- 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.
- **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.
- TRL actual 3/4.



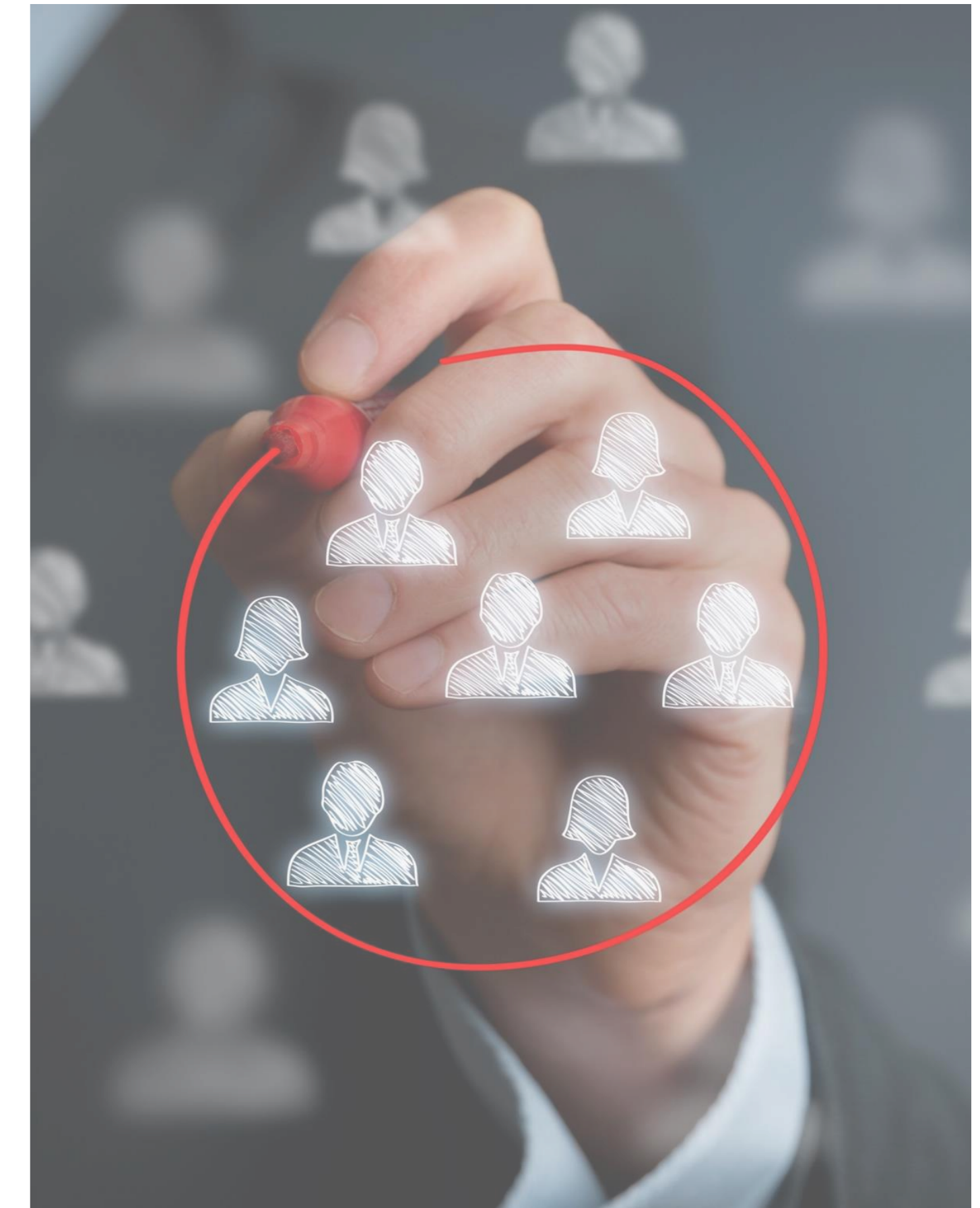
Oferta tecnológica: AnonimiX



Herramienta de anonimización basada en riesgo que **determina de forma automática la mejor forma de anonimizar un conjunto de datos**, maximizando su privacidad y utilidad:

- Soporta múltiples tipos de **datos estructurados**, incluyendo posiciones **GPS**.
- Ofrece diferentes algoritmos de anonimización.
- Dispone de **métricas de privacidad** que permiten medir el riesgo de reidentificación de los datos anonimizados, y **métricas de utilidad**, para medir su calidad.
- Soporta **Big Data** y anonimización en **tiempo real** (datos en *streaming*).
- Disponible mediante diferentes opciones de despliegue (herramienta *on-premise* o en la nube), y accesible a través de una interfaz web o API REST.

TRL actual: 7/8.



Oferta tecnológica: XDataHub



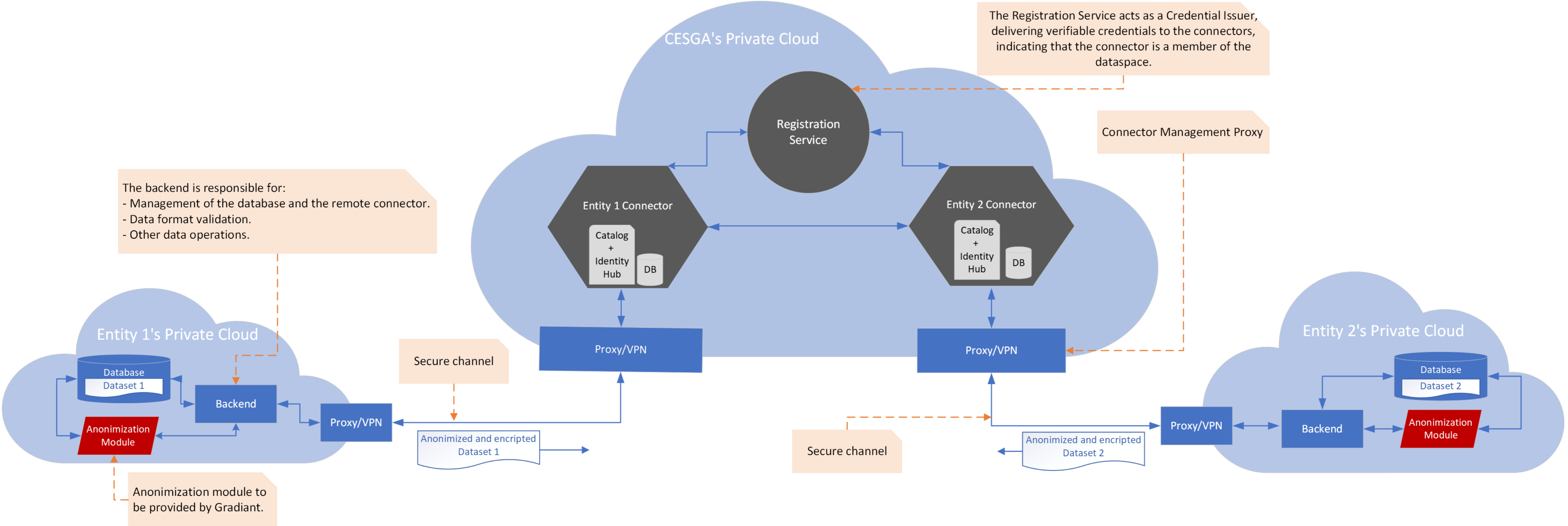
La anonimización transforma datos personales de forma que éstos no puedan ser vinculados con los individuos a los que pertenecen.

AnonymiX **simplifica y automatiza los procesos de anonimización en espacios de datos**, permitiendo:

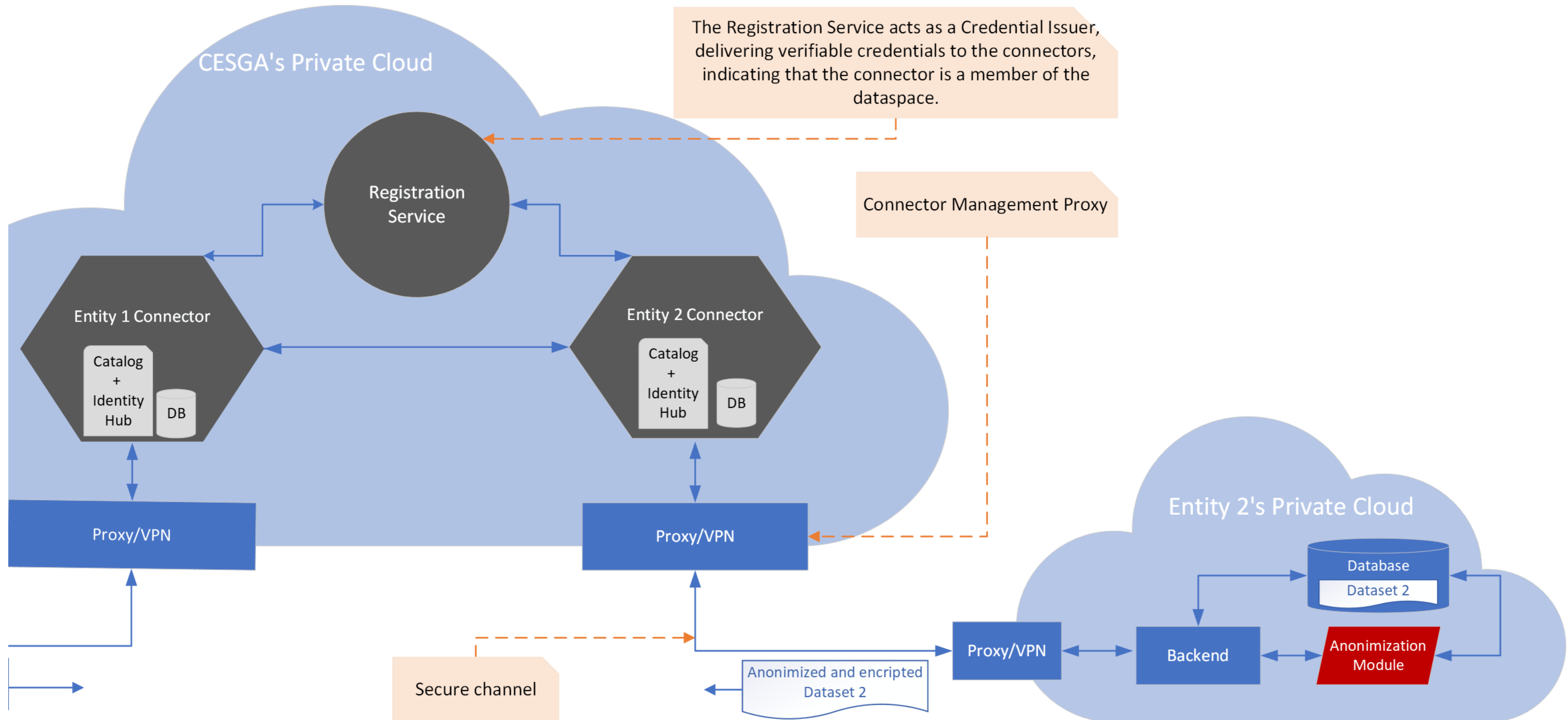
- **Compartir datos de forma segura** con terceros.
- Extraer valor de los datos, ya que la información anonimizada puede analizarse posteriormente.
- **Mejorar la protección de los datos sensibles**, evitando el acceso no autorizado y reduciendo el riesgo ante una brecha de seguridad.
- **Garantizar el cumplimiento regulatorio.**
- Incrementar la confianza, demostrando compromiso con la protección de los datos.



XDataHub y AnonimiX: posible despliegue (1/2)



XDataHub y AnonimiX: posible despliegue (2/2)





Muchas gracias por su atención

Helena Fernández, hfernandez@gradiant.org

(+34) 986 120 430 | gradiant@gradiant.org | www.gradiant.org





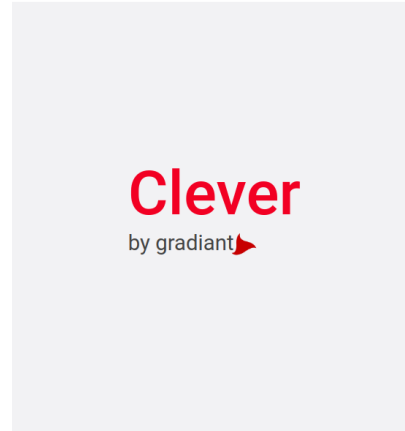
Clever Suite

Industria 4.0



¿Qué es Clever?

- Surge como una herramienta de apoyo a la llegada de la Industria 4.0 a las empresas manufactureras.
- **Finalidad:** Mejorar la operativa de los departamentos de la empresa para ayudar en la toma de decisiones en casos de uso concretos.
- Aspectos destacados:
 - Modular. Se adapta a las necesidades del cliente con diferentes funcionalidades. Algunas aplicables fuera del sector industrial
 - Dispone de la situación y funcionamiento actual de la empresa.
 - Funcionalidades avanzadas de IA y Data Analytics



Módulos funcionales

Planificación
de Producción

Optimización
de Compras

Ciclo de vida
de Producto

Monitorización

Data Analysis

Social Market



Optimización de Compras

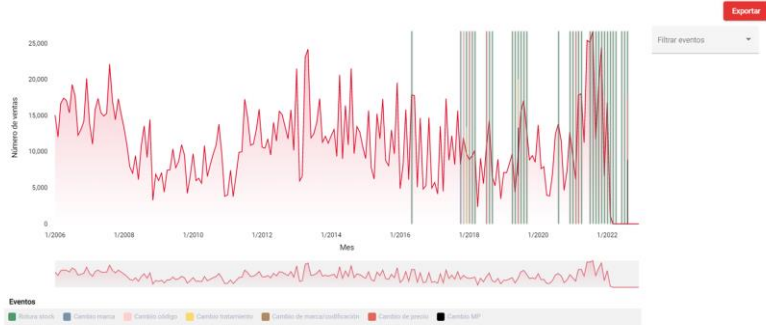


- Gestión de materias primas e indicadores cuantitativos y cualitativos
- Visualización de tendencias
- Pronóstico de indicadores de materias primas (ej. Precio)

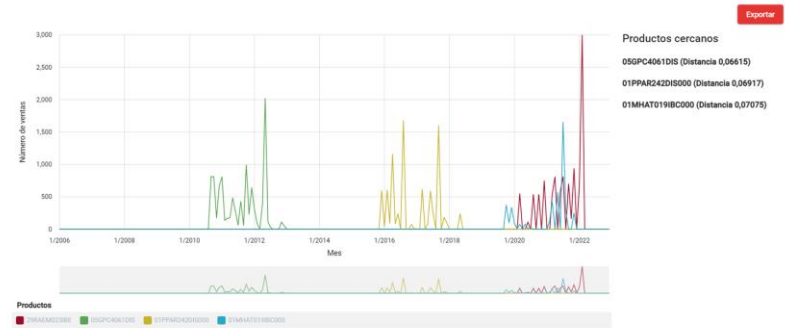


Ciclo de vida de productos

DETALLE DE PRODUCTO PARA 04REP021BC



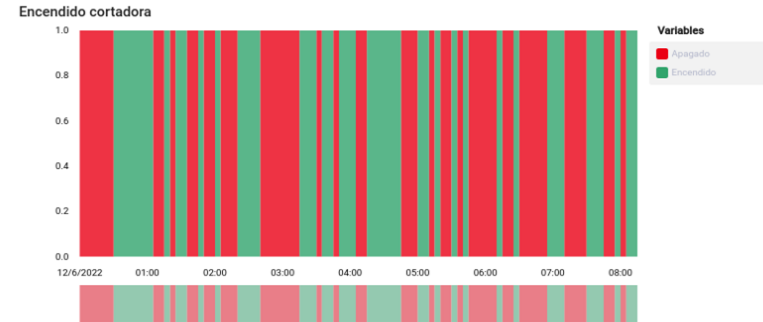
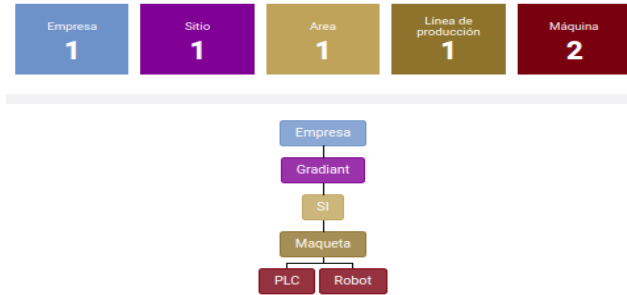
DETALLE DE PRODUCTOS CERCANOS PARA 29RAEM023IBE



- Gestión del ciclo de vida de productos y sus eventos:
 - Gestión de su historial de ventas y predicción futura de estas para nuevos productos.
 - Análisis de eventos (rotura de stocks, cambios de precio, etc.) aplicados sobre productos.
 - Cálculo de similitudes entre productos, en base al historial de ventas.



Monitorización

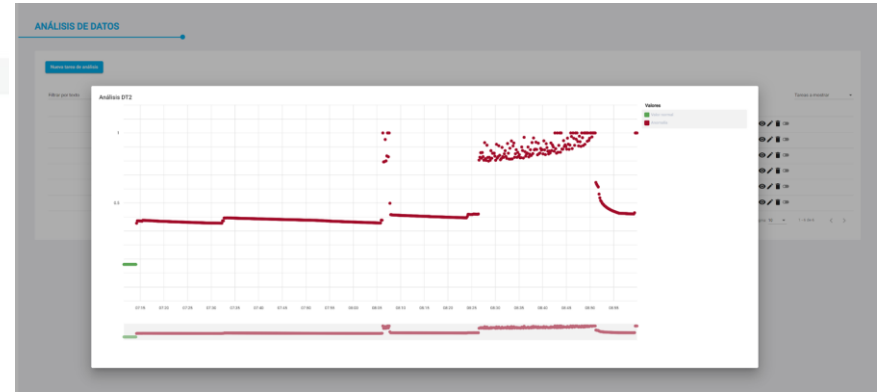


- Gestión jerárquica de los activos de la empresa y sus parámetros
- Monitorización de los activos (máquinas, sensores, bases de datos, etc.) mediante series temporales a través de diversos protocolos (OPC-UA, Modbus TCP, etc.)
- Creación de dashboards, análisis de productividad, generación de alarmas, seguimiento de KPIs, etc.



Data Analysis

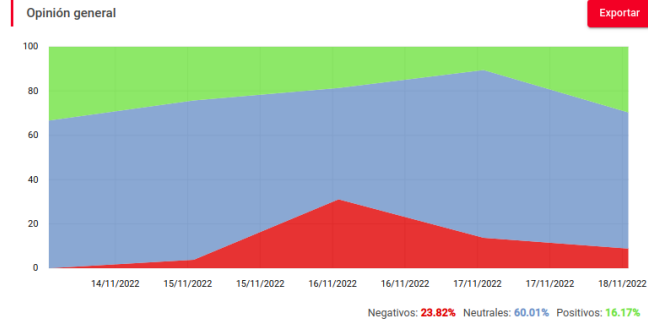
	Start Time	Duration	Run Name	User	Source	Version	Models	Parameters			Tags		
								batch_size	contamination	dropout_rate			
<input type="checkbox"/>	5 months ago	1.5min	-	ldiaz	train_ma...	355540	AutoEncode..._r1	32	0.1	0.2	AutoEncode...	['model_uef...	['a1485949...
<input type="checkbox"/>	5 months ago	1.5min	-	ldiaz	train_ma...	355540	AutoEncode..._r1	32	0.1	0.2	AutoEncode...	['model_uef...	['247974f...
<input type="checkbox"/>	5 months ago	12.8min	-	ldiaz	train_ma...	355540	-	-	-	-	-	-	-
<input type="checkbox"/>	5 months ago	12.7s	-	ldiaz	train_ma...	355540	ECCOMModel..._r4	-	0.1	-	ECCOMModel	['model_uef...	['247974f...
<input type="checkbox"/>	5 months ago	15.5s	-	ldiaz	train_ma...	355540	ECCOMModel..._r4	-	0.1	-	ECCOMModel	['model_uef...	['a1485949...
<input type="checkbox"/>	5 months ago	12.7min	-	ldiaz	train_ma...	355540	-	-	-	-	-	-	-
<input type="checkbox"/>	5 months ago	7.8s	-	ldiaz	train_ma...	355540	ECCOMModel..._r3	-	0.1	-	ECCOMModel	['model_uef...	['a1485949...
<input type="checkbox"/>	5 months ago	7.5s	-	ldiaz	train_ma...	355540	ECCOMModel..._r3	-	0.1	-	ECCOMModel	['model_uef...	['247974f...
<input type="checkbox"/>	5 months ago	8.0s	-	ldiaz	train_ma...	355540	ECCOMModel..._r2	-	0.1	-	ECCOMModel	['model_uef...	['spindle_b...
<input type="checkbox"/>	5 months ago	11.9s	-	ldiaz	train_ma...	355540	ECCOMModel..._r2	-	0.1	-	ECCOMModel	['model_uef...	['spindle_b...
<input type="checkbox"/>	5 months ago	10.5s	-	ldiaz	train_ma...	355540	ECCOMModel..._r1	-	0.1	-	ECCOMModel	['model_uef...	['spindle_b...
<input type="checkbox"/>	5 months ago	7.7s	-	ldiaz	train_ma...	355540	ECCOMModel..._r1	-	0.1	-	ECCOMModel	['model_uef...	['spindle_b...
<input type="checkbox"/>	5 months ago	5.3min	-	orgonatez	-	159e4e	AutoEncode..._r2	32	0.1	0.2	AutoEncode...	['model_uef...	['spindle_r...
<input type="checkbox"/>	5 months ago	11.3s	-	ldiaz	train_ma...	e876c2	ECCOMModel8	-	0.1	-	ECCOMModel	['model_uef...	['spindle_r...



- Ejecución de análisis y experimentos (batch o streaming) sobre los datos del módulo de monitorización mediante algoritmos simples de ML/IA (ej. Correlaciones, predicción, detección de anomalías...)
- Integración con Jupyter Hub para realizar análisis personalizados



Social Market



216.6k	+46,54%	5.5M	11,06%	35,05%
Usuario	Incremento de la audiencia	Alcance potencial	Tasa de amplificación	Incremento del alcance

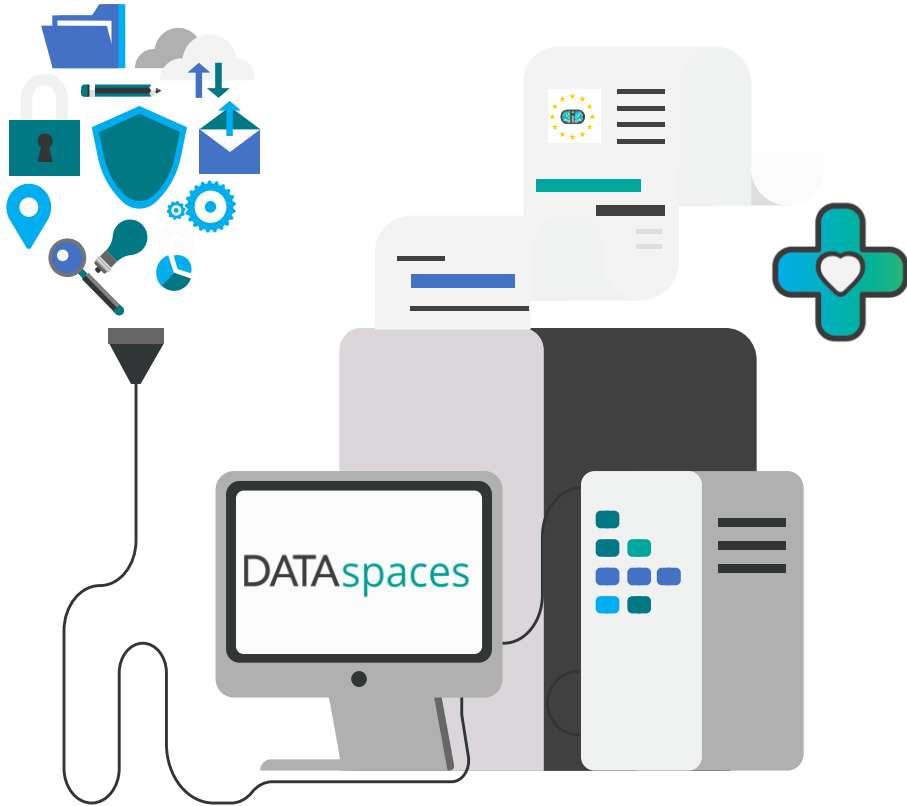
Usuario	Opiniones	Alcance potencial		
1	Kaja Boshman	33716	23.2%	★★★★☆
2	Rahven Alkewerth	1228	23.7%	★★★★☆
3	Cecile Stokes	6343	6.8%	★★★★☆

Usuario	Opiniones	Positivas	Negativas	
1	Dajana Leeman	136	33	██████████
2	M. Ornel Manó	174	28	██████████
3	André Braker	136	2	██████████

- Analizar y visualizar el impacto de la empresa y sus productos en RRSS
 - Usuario genera un contexto para la búsqueda en las redes: RSS, Instagram o Twitter (X, API de pago) o Reddit.
 - Análisis de tendencias. Agrupa contenidos por temática.
 - *Polarity Analysis*: Analiza sentimientos (positivo, negativo o neutro), según su texto.
 - *Interest Groups Analysis*: Etiqueta grupos del contexto según la temática del texto.



Gaia-X salud en Galicia



3 de abril



Biopolo Sionlla
CSIEB

Organizan:



Demostrador DATALIFE



- eDIH Datalife
- Programa CCMM → Plataforma Integrada Datos Mariños (10 M€)
- Necesidades investigadores e empresas
- Ecosistema biosanitario
- PERTE Chip (47,4 M€) → Datacenter
- Estratexia CESGA
→ Investigación
→ Posicionamento en Datos



GAIA X

ESPACIOS DE DATOS DE SALUD

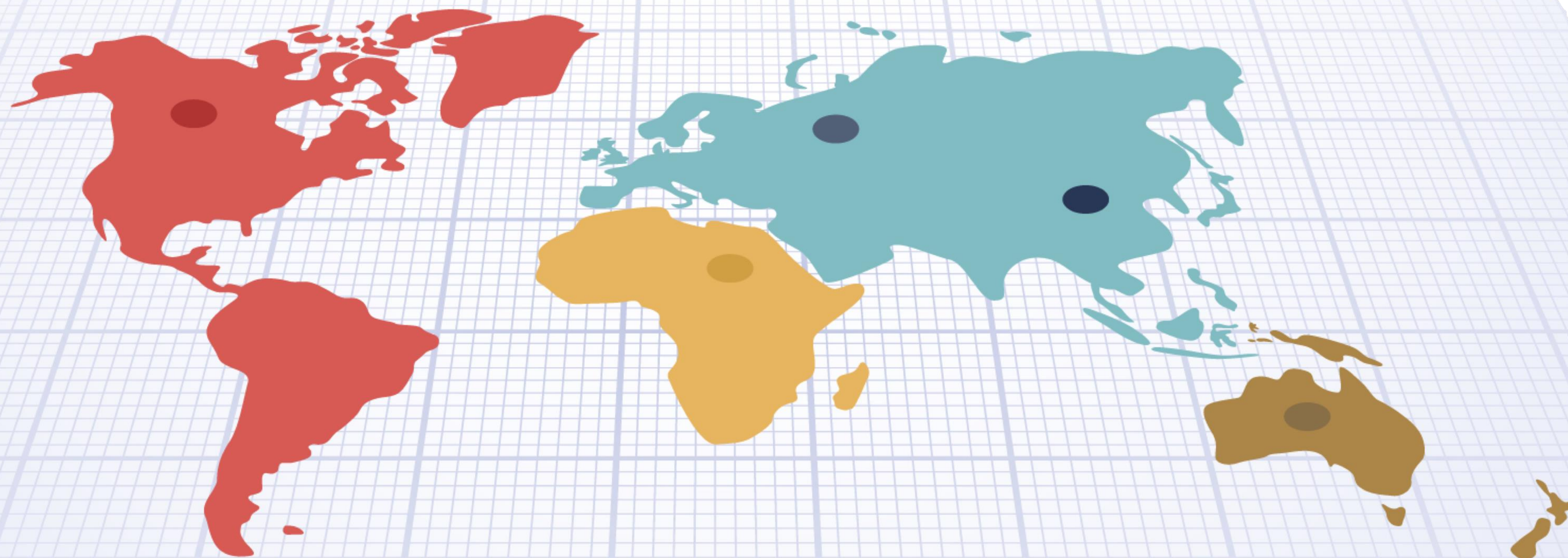
April 2024





THE CHALLENGE: ACCESS TO DATA

Shortage of labeled data has been holding the surge of deep learning in healthcare back, as sample sizes are often small, patient information cannot be shared openly, and multi-center collaborative studies are a burden to set up.



SOLUTION: FEDERATED LEARNING

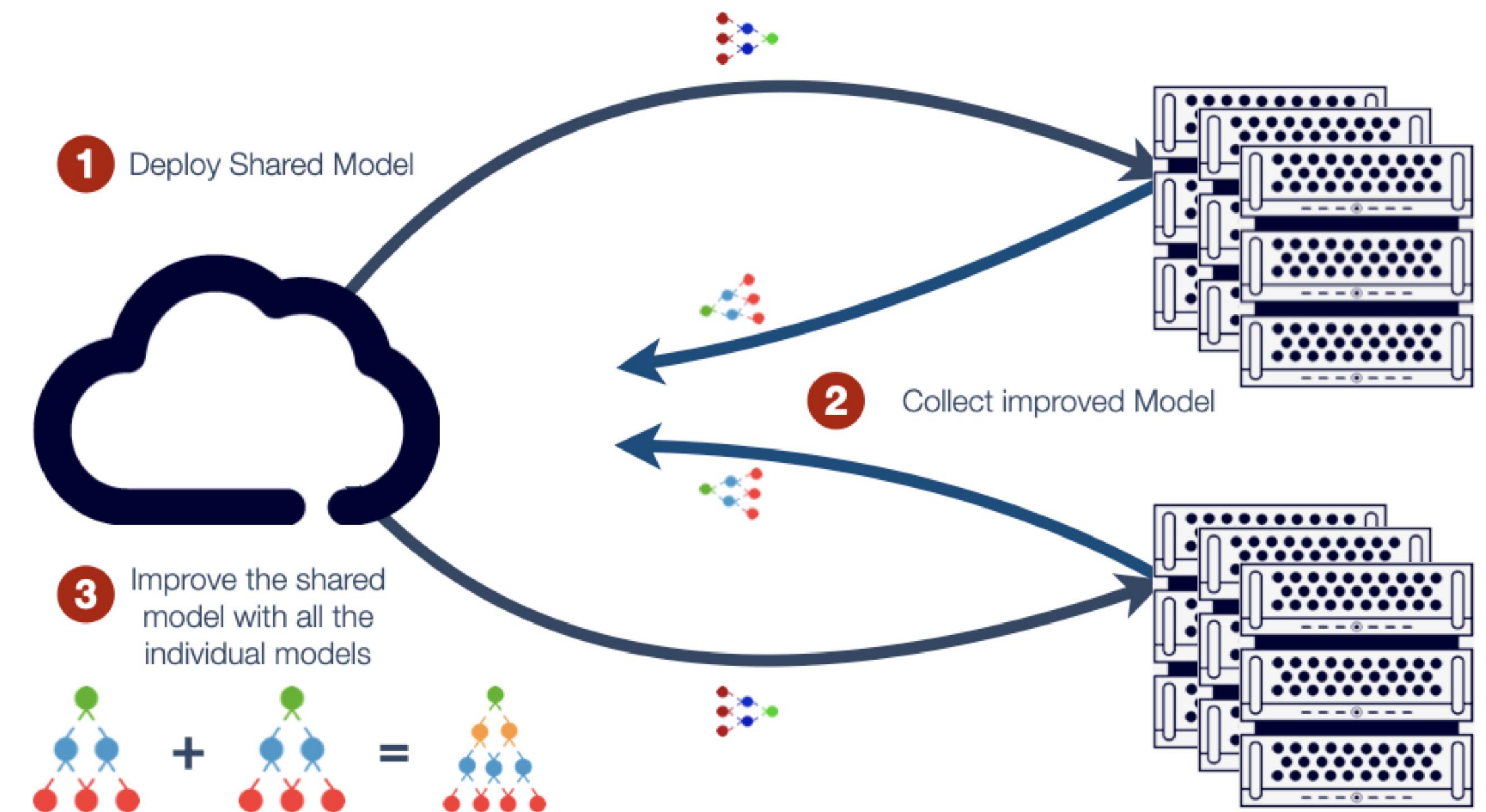
COMPUTATIONAL PRIVACY.

Don't move data, move the models to the data.

Privacy Enhancing Technologies (PETs) enable businesses to learn together while people's personal data remains private and confidential.

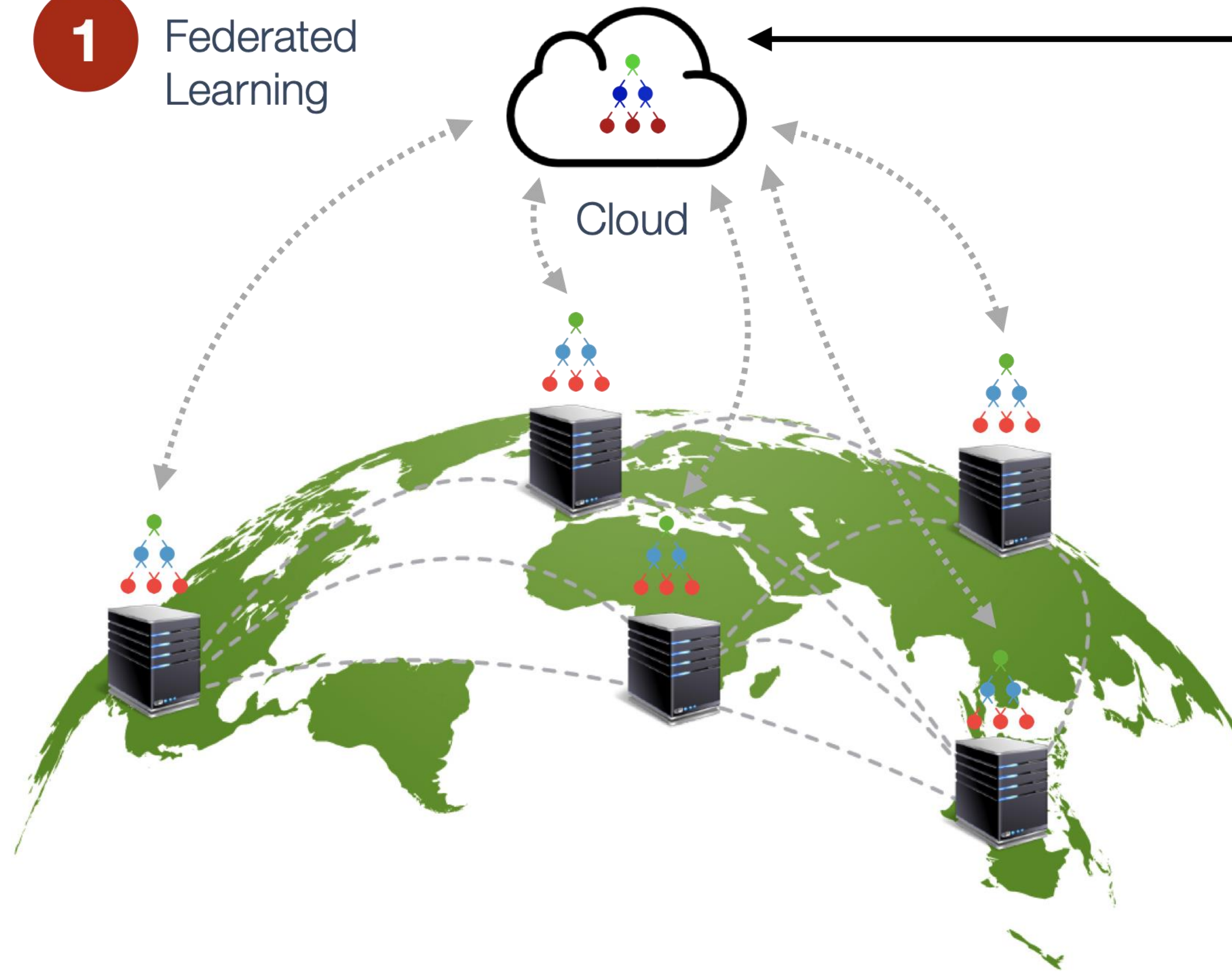
Federated Learning finds useful patterns in historical data in order to make predictions while ensuring people's individual data doesn't move.

Partners keep control and confidentiality of data, while complying with privacy regulations.

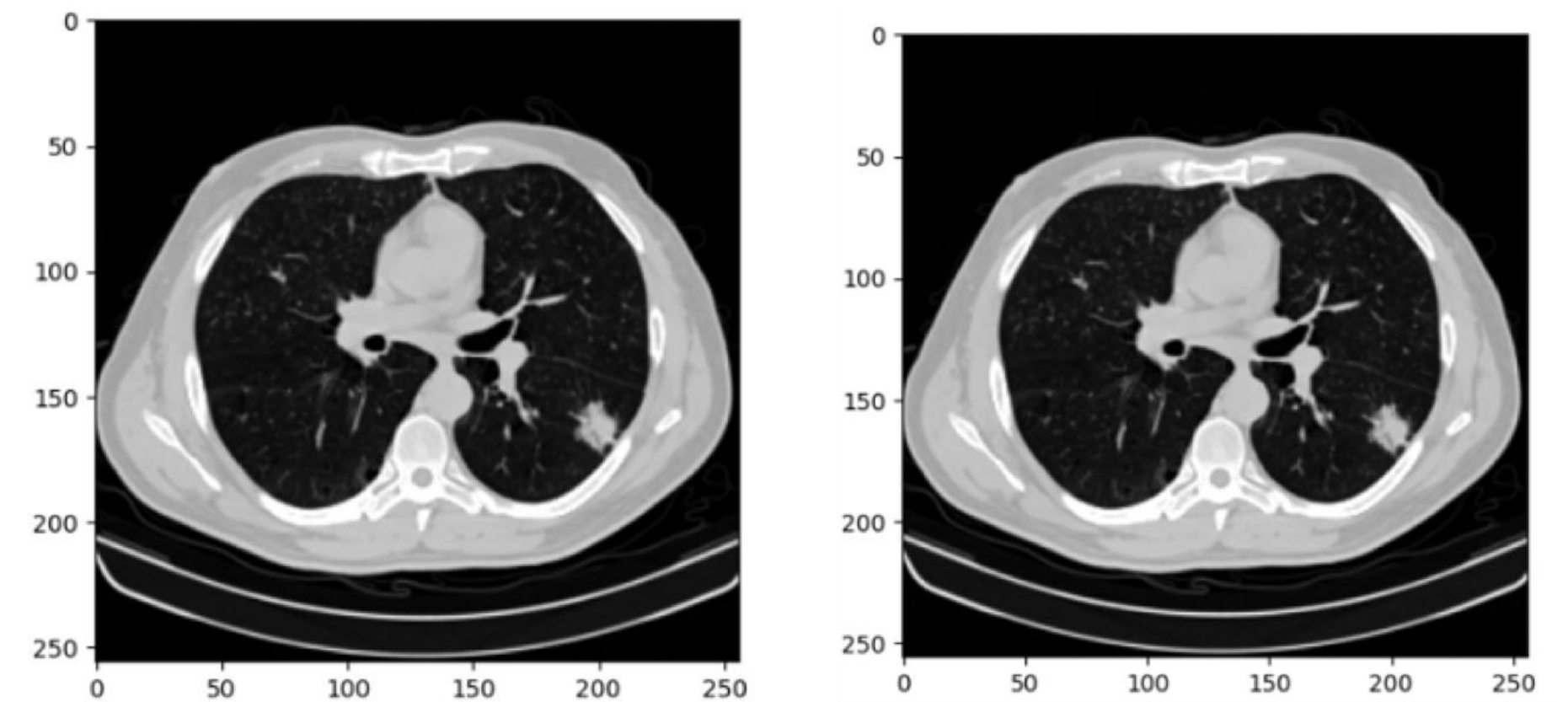


USE CASE: ANALYZE CT-SCANS

1 Federated Learning



2 API to Analyze CT Scans



DEMO ACURATIO PLATFORM

USE CASES: EXAMPLES

HEALTHCARE USE CASES

- Medical Image Analysis.
 - Drug Discovery.
 - Pharma collaborations.
 - Clinical trials: analytics and consolidations.
 - Private record linkage for Analytics.
 - Copilots for every specialty.
 - Federated recommender system for medical codes (CIE 11 and DSM-5).
 - Chat-gpt from unstructured data.
- ...

AI is a need due to the shortage of healthcare professionals.



Share knowledge, without sharing data

