

tecna:a

MEMBER OF BASQUE RESEARCH
& TECHNOLOGY ALLIANCE

gipuzkoatic)))

Computación Segura

Seguridad de la IA y la
Economía del Dato



CAMARA DE GIPUZKOA
GIPUZKOAKO BAZKUNDEA

**IA es una
Herramienta
Necesaria**





**Pero la IA
también puede
ser atacada**



Poisoning Attacks

Ataques de envenenamiento


Tipos:

- **Data Poisoning:** introduce información falsa o alterada en el conjunto de datos, haciendo que el modelo aprenda de forma incorrecta y tome decisiones erróneas.
- **Model Poisoning:** manipulación directa del modelo de IA para introducir vulnerabilidades o comportamientos no deseados.

Impacto:

- **Degradación del Rendimiento:** clasificación incorrecta de datos, afectando a su precisión y fiabilidad.
- **Pérdida de Control:** atacantes pueden causar que el modelo actúe de manera impredecible o contra el interés de su propietario,
- **Daño a la Reputación:** un modelo comprometido puede dañar la reputación de la organización que lo ejecuta.





**No Podemos Aislar
la IA y sus datos
de entrenamiento**

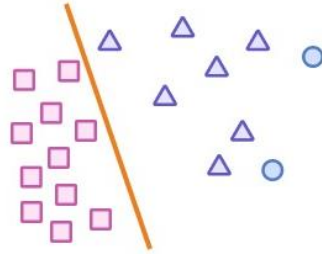


**La IA necesita
Datos para
aprender**

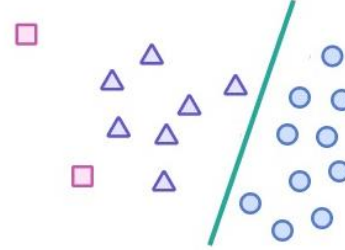


**Billetes que
salen de la pared**

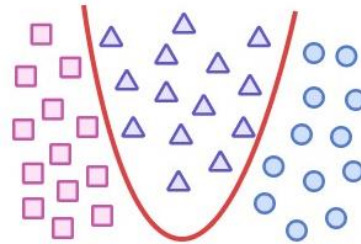
Classifier-1



Classifier-2



Ensemble Classifier



Data Economy



Compartir información para
obtener mejores insights es
el futuro...

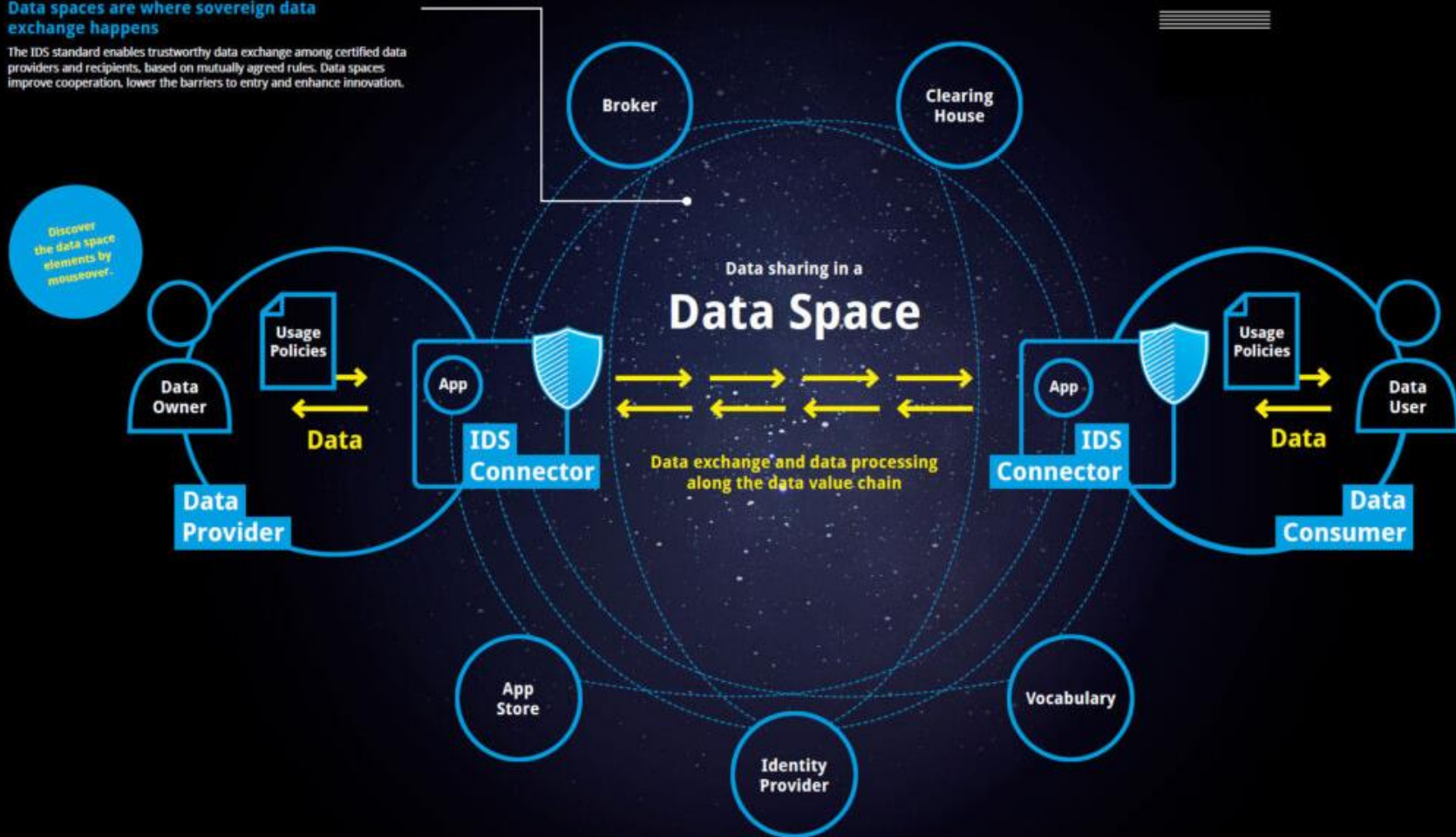


Pero una vez que compartimos la información perdemos el control sobre la misma



Data spaces are where sovereign data exchange happens

The IDS standard enables trustworthy data exchange among certified data providers and recipients, based on mutually agreed rules. Data spaces improve cooperation, lower the barriers to entry and enhance innovation.



Contributions of IDSA to the success of GAIA-X



The IDS reference architecture serves as an initial impulse for GAIA-X. As a basis for an open ecosystem, it enables provider and consumer of data to connect in a secure, interoperable and sovereign way. Combined with highly available storage and efficient processing of data GAIA-X has the potential to create a secure and trustworthy data infrastructure based on European values.



CONCEPTUAL



- Concept for data sovereignty
- Mindset for the value of data sharing

TECHNICAL



- A trustworthy architecture for data sovereignty
- Reference Architecture V3.0

STANDARD



- Formal DIN standard on its way to ISO
- security gateway for sharing data and services

EU DATA STRATEGY



- Design principles and functional building blocks for data spaces

COMMUNITY OF PRACTICE

50+

- Concrete implementations with high TRL or even commercial offering
- National and European research projects with 100+ Mio. € volume

GLOBAL REACH

100+



- 100+ members from 20 countries all over the world
- Companies, research organizations, industry associations

IDSA HUBS



- Formal agreements with organizations in 9 countries to proliferate data sovereignty on country level

CERTIFICATION



- Formal certification states coherence to reference architecture, interoperability and software quality

GLOBAL LIAISONS



- Harmonization of reference architecture with leading global initiatives

European Strategy for Data

A common European data space, a single market for data



Common European data spaces

Rich pool of data
(varying degree of
accessibility)

Free flow of data
across sectors and
countries

Full respect of GDPR

Horizontal
framework for data
governance and data
access



Health



Industrial &
Manufacturing



Agriculture



Finance



Mobility



Green Deal



Energy



Public
Administration



Skills

- Technical tools for data pooling and sharing
- Standards & interoperability (technical, semantic)
- Sectoral Data Governance (contracts, licenses, access rights, usage rights)
- IT capacity, including cloud storage, processing and services

The existing Split X-Model

Disjoint Data & Infrastructure Ecosystems



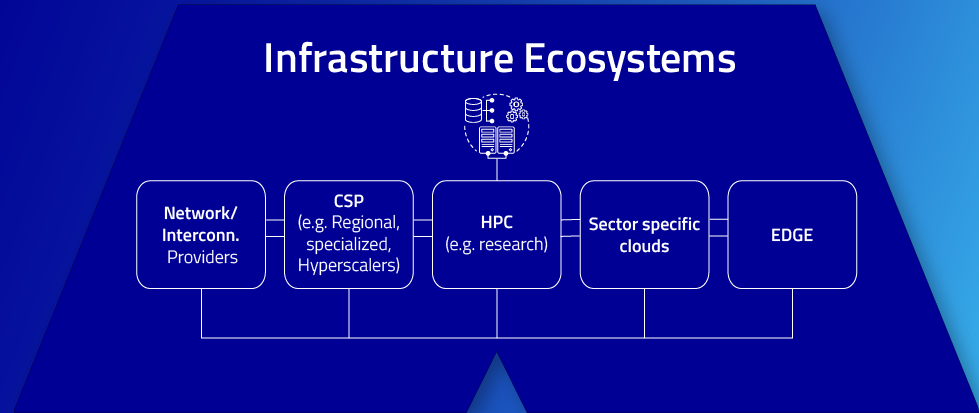
Data

Untapped, fragmented, disjoint, no secure exchange mechanism



Infrastructures

Segregated, non-reversible, non-interoperable, closed architectures, private standards



Our X-Model

Connecting Data & Infrastructure Ecosystems



Advanced Services

New (Cross-) Sector Innovations / Applications build from service composition.



Data Spaces / Federations

Interoperable & portable (Cross-) Sector data-sets and services.



Data Exchange

Anchored contract rules for access and data usage.



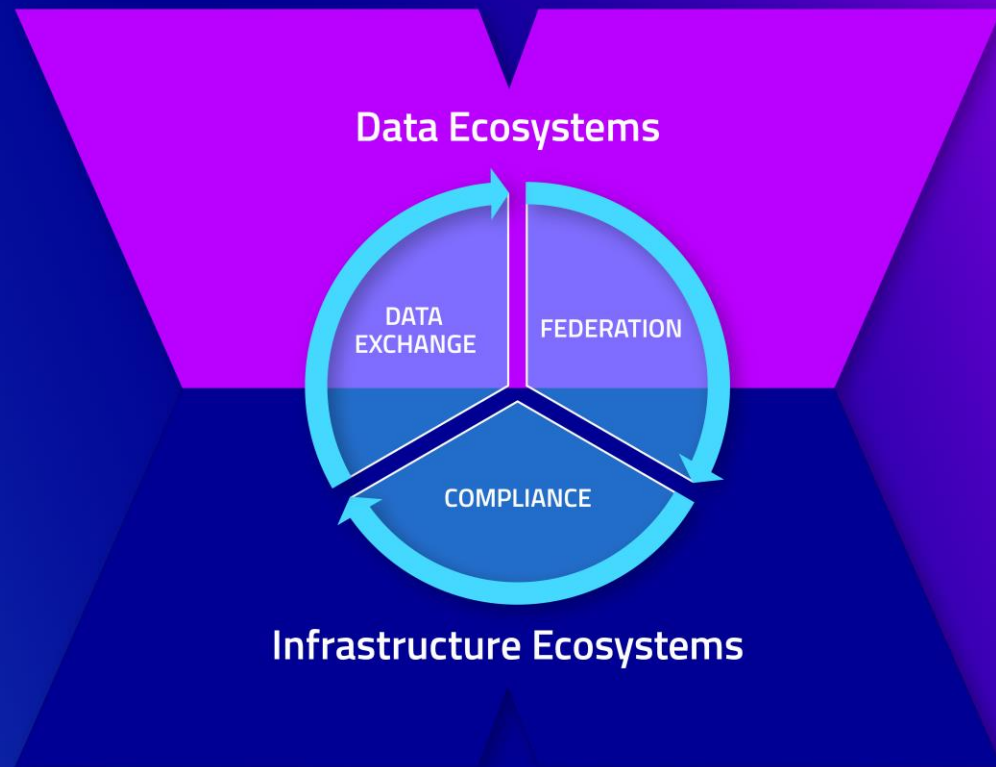
Gaia-X Compliance


Decentralized services to enable objective and measurable trust.



Label framework

Gaia-X and ecosystem specific Labels to ease market adoption through autonomy and self-determination.

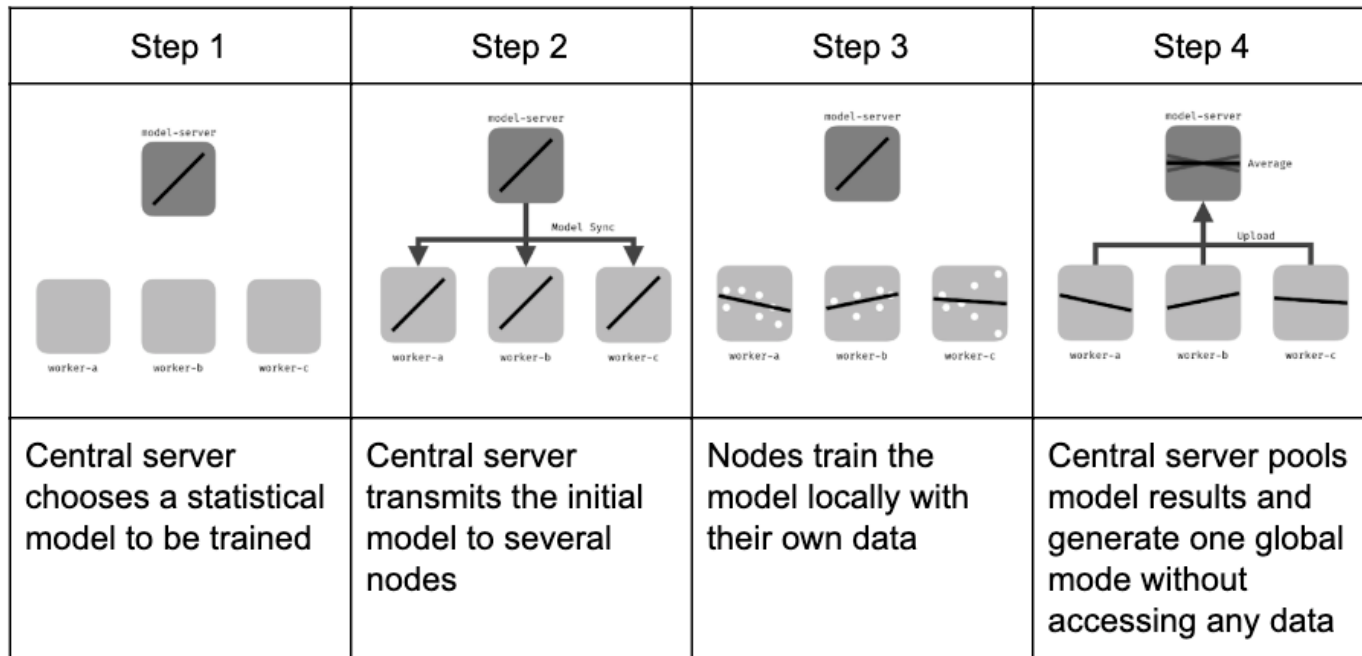




Riesgos de los Espacios de Datos (Actuales)

Cryptographic vs Policy-
Based Enforcement

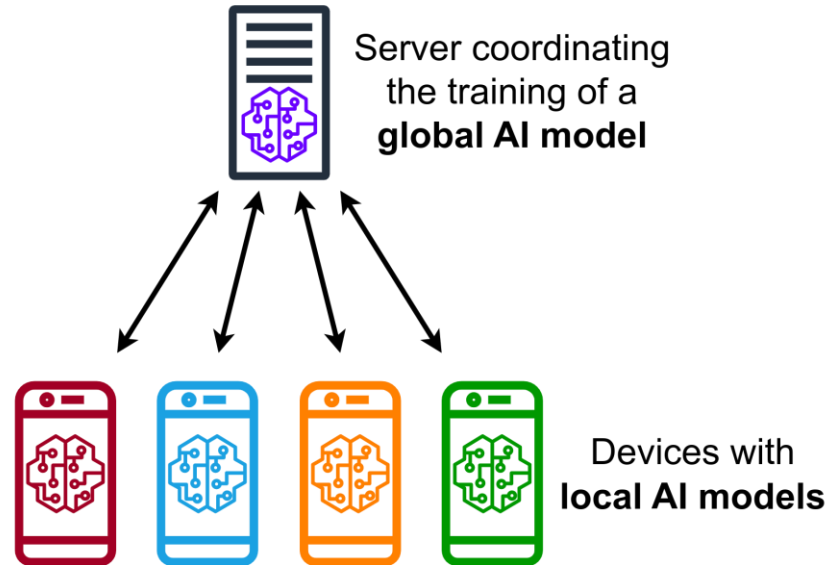
Federated Learning



Riesgos Federated Learning

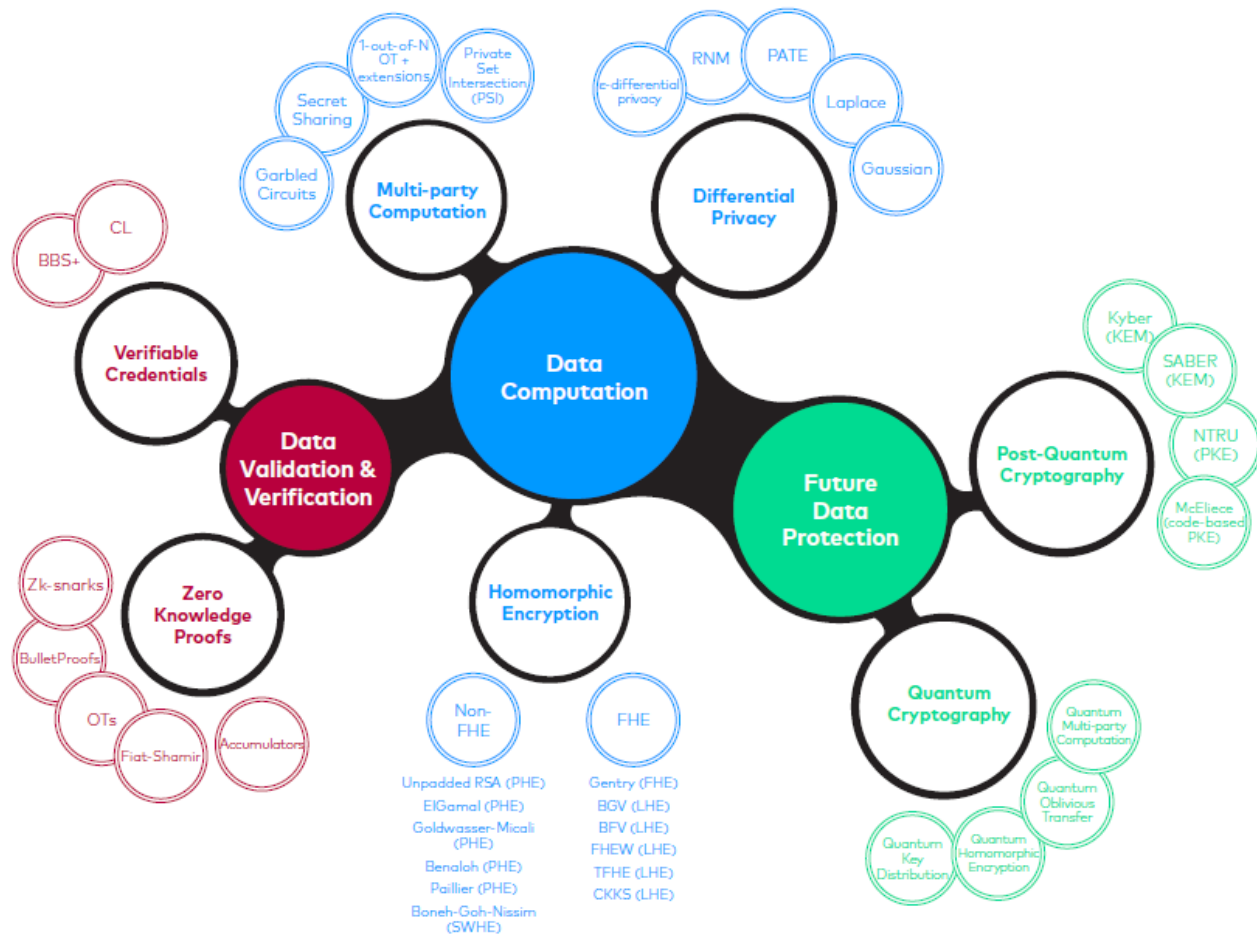
Tipos:

- **Reconstrucción:** buscan reconstruir datos de entrada originales a partir de los gradientes o actualizaciones de modelos compartidos durante el entrenamiento. El atacante se aproxima a los datos que contribuyeron al modelo parcial/global.
- **Inferencia de Propiedades:** deducir atributos o características específicas de los participantes del modelo, como enfermedades en un conjunto de datos de salud, sin reconstruir los datos completos
- **Inferencia de Membresía:** Determinan si un registro de datos específico fue utilizado en el entrenamiento. Cuidado con datos sensibles.



Necesitamos **compartir** los datos **de forma**
segura: **protegiendo tanto** los **datos** como
los **algoritmos** o consultas realizadas





Privacy Preserving Computing





MULTIPARTY COMPUTATION

Criptografía HOMOMÓRFICA





Tres principales casos de uso

Estudios o análisis estadísticos sobre datos de terceros (ej. salario medio de los empleados de un sector)





Generar nuevo conocimiento o
Entrenamiento de algoritmos
(protegiendo nuestro algoritmo y sus datos)

Ofrecer un servicio sobre datos de terceros basado en nuestro conocimiento o algoritmo

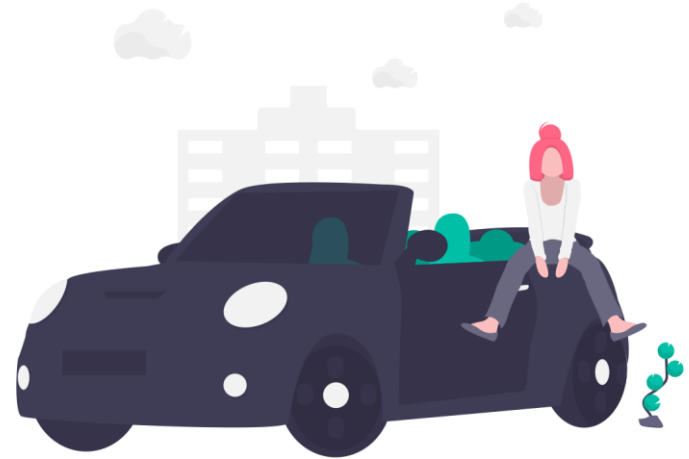


Industria

- **Mantenimiento Predictivo** como servicio
- **Mejora de productos** y **detección de fallos** en base a info de explotación
- Entrenamiento **Gemelos Digitales** y su aplicación para la identificación de anomalías

Automoción

- Algoritmos **Conducción autónoma**
- Mejora de **infraestructuras** en base a información proporcionada por los vehículos (baches, puntos negros, etc.).



Salud

- Investigación y **Ensayos clínicos**
Servicios diagnósticos basados en Imagen, ADN, etc.
- Mejorar la **interoperabilidad** entre sistemas de salud (“queries respetuosas”)



Conclusiones

- La **IA** es una **Herramienta Necesaria** para Nuestro Futuro.
- La **Importancia del Acceso a los Datos**: para que la IA sea precisa necesita acceso a grandes cantidades de datos.
- La **UE está impulsando Iniciativas** y **Regulaciones** para impulsar el intercambio de datos, no exentas de **ciertos riesgos** en sus primeras versiones.
- La **Criptografía como Habilitadora** de la compartición segura del dato, impulsando su adopción en el futuro.

Recomendaciones

- **Aumentar la Conciencia** sobre los posibles **riesgos asociados** con la IA y el intercambio de datos.
- **Promover el Intercambio de Datos Responsable** para prevenir sesgos, mejorar las predicciones y proteger la privacidad del usuario.
- **Aprovechar las Técnicas de Computación Segura** para preservar la privacidad/confidencialidad de los datos compartidos.



MEMBER OF BASQUE RESEARCH
& TECHNOLOGY ALLIANCE



CAMARA DE GIPUZKOA
GIPUZKOAKO BAZKUNDEA

Oscar Lage

oscar.lage@tecnalia.com

@Oscar_Lage



tecnalia.com