ORGANISED BY

BDV BIG DATA VALUE ASSOCIATION

IN COLLABORATION WITH

IFE

NTNU

SINTEF

WESTERN NORWAY RESEARCH INSTITUTE
VESTLANDSFORSKING

DATAWEEK

JOIN.LEARN.SHARE.GET VALUE

DATA FJORDS: UNLOCKING AI FOR INDUSTRY & SOCIETY

5-6 May 2026 / Oslo

SUPPORTED BY

The Research Council of Norway

data-week.eu
#DataWeek26

DATAWEEK

JOIN.LEARN.SHARE.GET VALUE

Data Labs: their role in Europe's Data and AI Landscape

ORGANISED BY BDV BIG DATA VALUE ASSOCIATION

SUPPORTED BY The Research Council of Norway

IN COLLABORATION WITH IFE NTNU SINTEF WESTERN NORWAY RESEARCH INSTITUTE VESTLANDSFORSKING

Objectives of the session

- Clarify the role of AI Data Labs in the European landscape
- Discuss about their complementarities and overlaps with existing initiatives
- Identify strategic directions and recommendations for the future development of Data Labs

DATAWEEK

JOIN.LEARN.SHARE.GET VALUE

DATA LABS: THEIR ROLE IN EUROPE'S DATA AND AI LANDSCAPE

Daniel Saez
ITI, BDVA BoD

Heidi Laine
CSC – IT Center for Science, LUMI AIF

Ville Tenhunen
EGI

Jürgen Vangeyte
ILVO, CEADS

Sergio Gusmerolli
Politecnico di Milano

Jorge Amor Río
Gradiant, 1Health AIF

Mark Dietrich
SAGE Dataspace

Jan Martinovic
IT4Innovation, CZAI

Agenda

DATA LABS: THEIR ROLE IN EUROPE'S DATA AND AI LANDSCAPE

Presentations

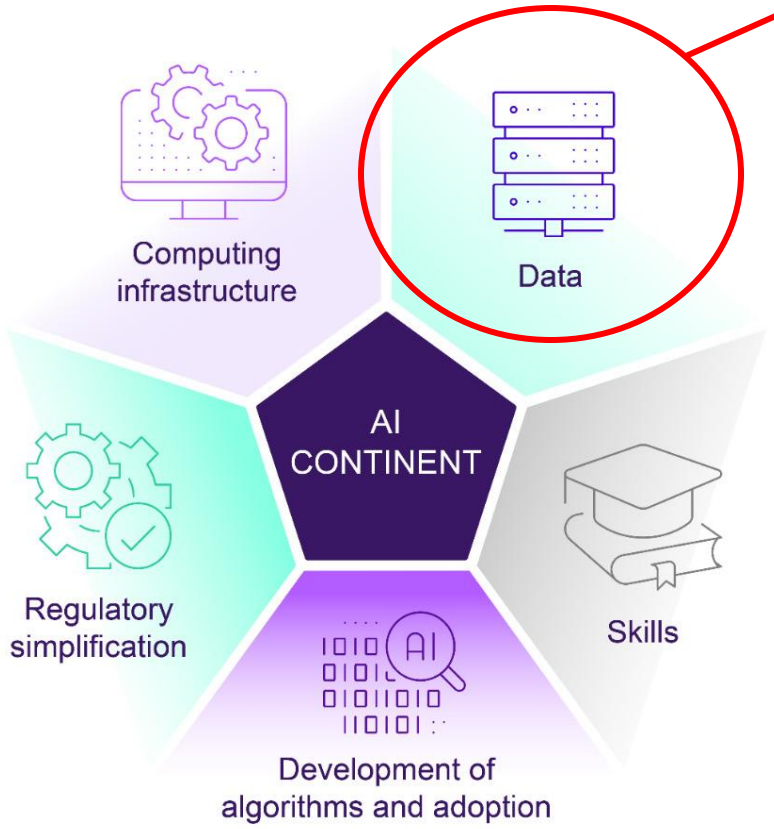
- Introduction and motivation (Daniel Alonso, BDVA)
- BDVA initiative on Data Labs (Daniel Saez, ITI, BDVA BoD)
- AI Factory & Data Lab perspective (LUMI AIF and Data Lab on Scientific Data, Heidi Laine, CSC-IT)
- Data exploitation platform: tools and opportunities for Data Labs (RI-SCALE, Ville Tenhunen, EGI)
- Data Spaces – Data Labs – AI Factories
  A match made in heaven! (Jürgen Vangeyte, ILVO)
- View from EDIHs (Sergio Gusmeroli, Politecnico de Milano)

Panel discussion

- Moderator: Daniel Saez (ITI, BDVA BoD)
- Panelists:
  - Heidi Laine (LUMI AIF, Data Lab)
  - Ville Tenhunen (EGI)
  - Jorge Amor Río (Gradiant, 1Health AIF)
  - Jan Martinovic (IT4Innovation, CZAI)
  - Mark Dietrich (Bloodstone, SAGE)

AI Continent Action Plan

(9 April 2025)



Data Labs

- Part of the AI Factories initiative
- Turn fragmented data sources into a trusted and accessible resource for AI development
- Bring together and federate data from different AI Factories covering the same sectors
- Link to the corresponding Common European Data Spaces
- Make data available to AI developers under appropriate conditions
- Cleaning and enriching
- Other technical tools (e.g. standardised formats, synthetic data, shared technical building blocks)
- Foster interoperability
- Data-pooling services
- **Further developed in the European Data Union Strategy**

Towards a European AI-data value Ecosystem (05.06.2025):
https://bdva.eu/news/towards-a-european-ai-data-value-ecosystem/

European Data Union Strategy *(19 November 2025)*

Data Spaces and Data Labs: The Building Blocks of Europe's AI Ecosystem"

Data spaces -> trusted infrastructures where data is governed and made available

Data Labs -> operational interface that enables its safe, value-adding use for AI

- **First data labs** established under the AI Factories through EuroHPC
- **Further Data Labs** to be set up independently in other domains to address specific sectoral or research needs
- Work closely with EDIHs, which act as user-facing contact points and help match data needs with concrete applications

Data at the core of Europe's Digital Strategy (18.07.2025): https://bdva.eu/news/data-at-the-core-of-europes-digital-strategy/

BDVA Initiative on Data Labs

Daniel Sáez-Domingo (ITI)
Daniel Alonso Román (BDVA)

ORGANISED BY BDV BIG DATA VALUE ASSOCIATION
SUPPORTED BY The Research Council of Norway
IN COLLABORATION WITH IFE NTNU SINTEF WESTERN NORWAY RESEARCH INSTITUTE VESTLANDSFORSKING

BDVA Data Labs initiative

- BDVA member of the EuroHPC JU, active engaged across Data Spaces, AI and data technologies, governance, trust, and sectoral applications
- BDVA at the intersection of AI Factories, data ecosystems, and AI technologies.
- **BDVA uniquely suited to coordinate, support and guide Data Labs-related activities** in a coherent and cross-cutting manner

Provide a clear, visible, and coordinated framework within BDVA for all activities related to Data Labs, while fully leveraging existing Task Forces, work and initiatives

Ensure ***coherence and alignment*** across Data Labs initiatives and domains

Provide a ***focal point for engagement*** with external stakeholders and policymakers

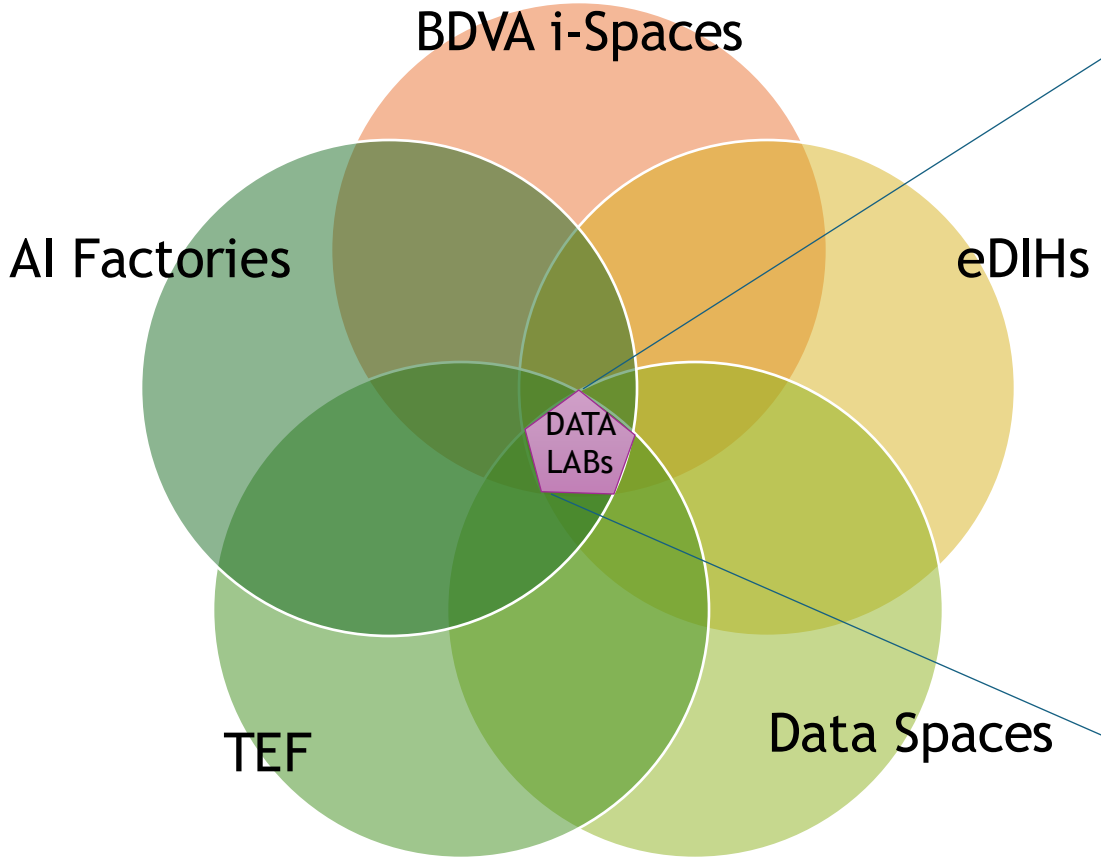
Support evidence-based ***guidance for future phases*** and investments

Strengthen BDVA's strategic ***contribution to European data and AI policies***, including the next Multiannual Financial Framework (MFF)

First step: Ecosystem mapping and clarification

AI Continent Action Plan
(9 April 2025)

European Data Union Strategy
(19 November 2025)



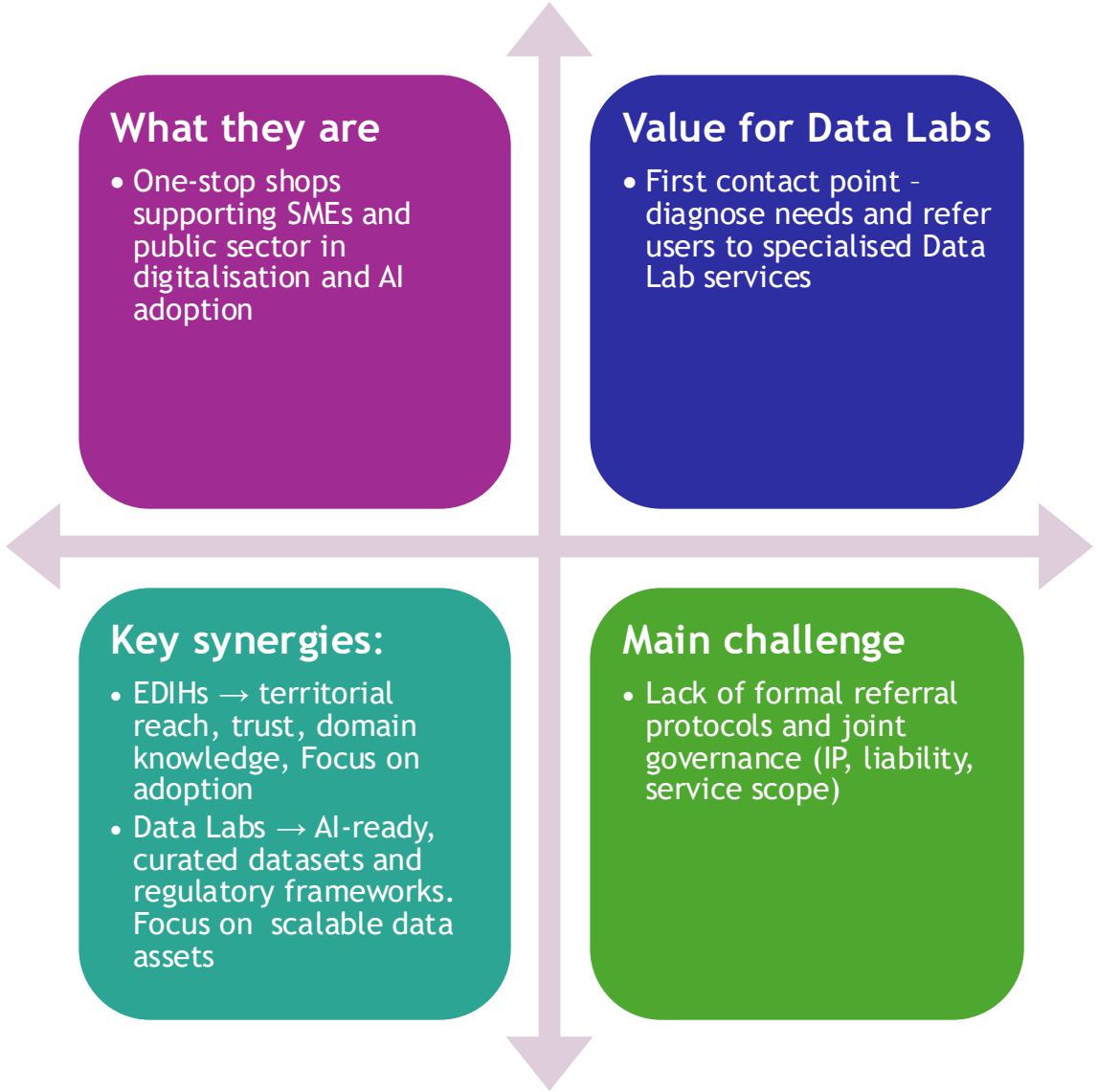
DATA LAB

Neutral and trusted entity	Addressing limited availability of **data for AI training**	Specialised facilities, **linking** data & AI ecosystems
Filling high Security requirements	Is a trusted party for data spaces and for PPAA	Technological capabilities for data management, curation, data generation, ...
No interest in exploiting data, but in preparing them for others	Training and regulation orientations	Interest data identification/ scouting
Annonymization, pseudonnymization, synthetic data	Data labeling, cleaning, ...	Sustainable business model

AI tools for data quality verification of those data generated for training models

Data Labs & eDIHs

Entry points to Data Labs



Data Labs & TEFs

From experimentation to reusable data

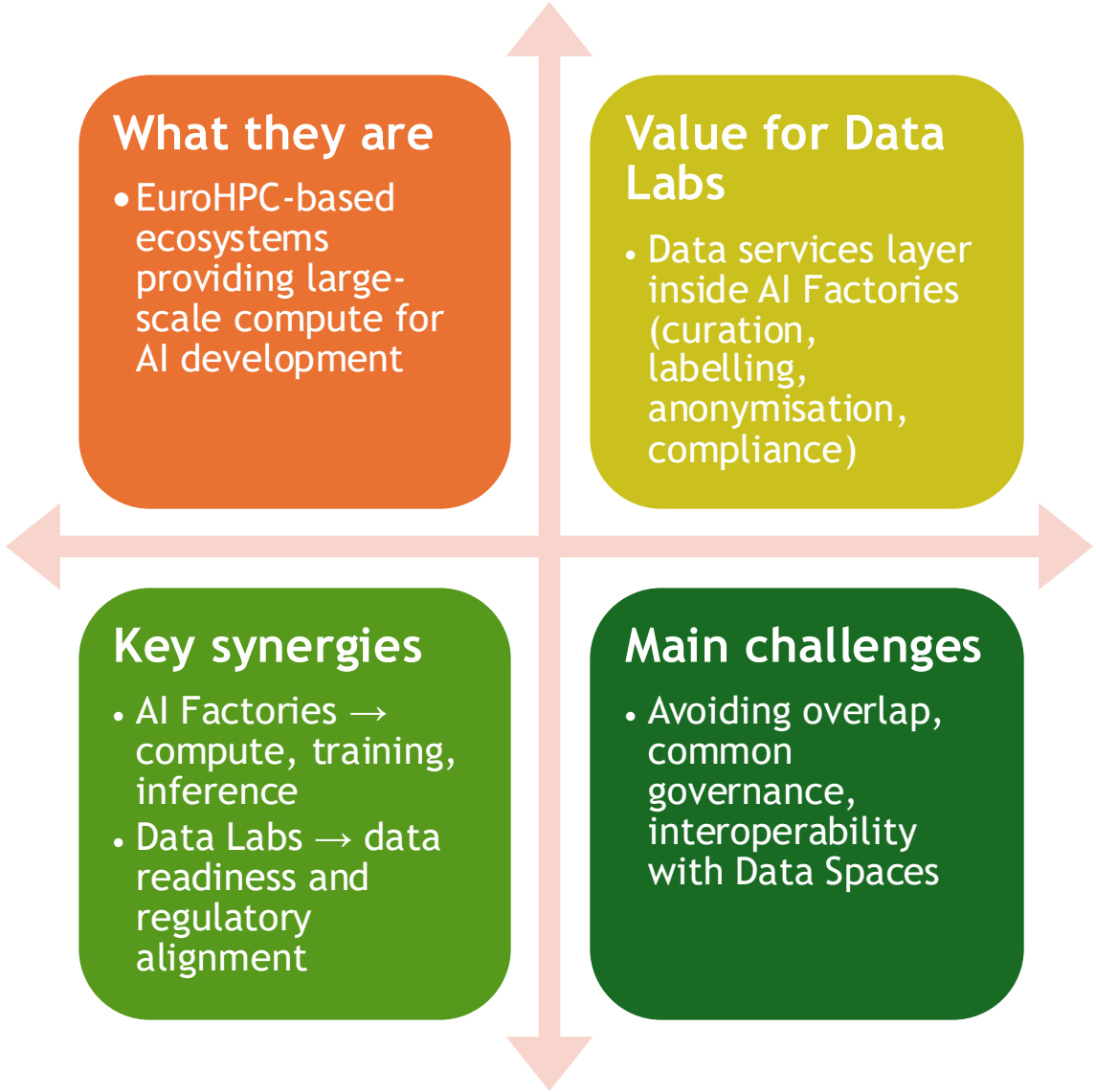
Data Labs & Data Spaces

The governed data backbone

Data Labs & AI Factories

First formal environment where Data Labs are instantiated

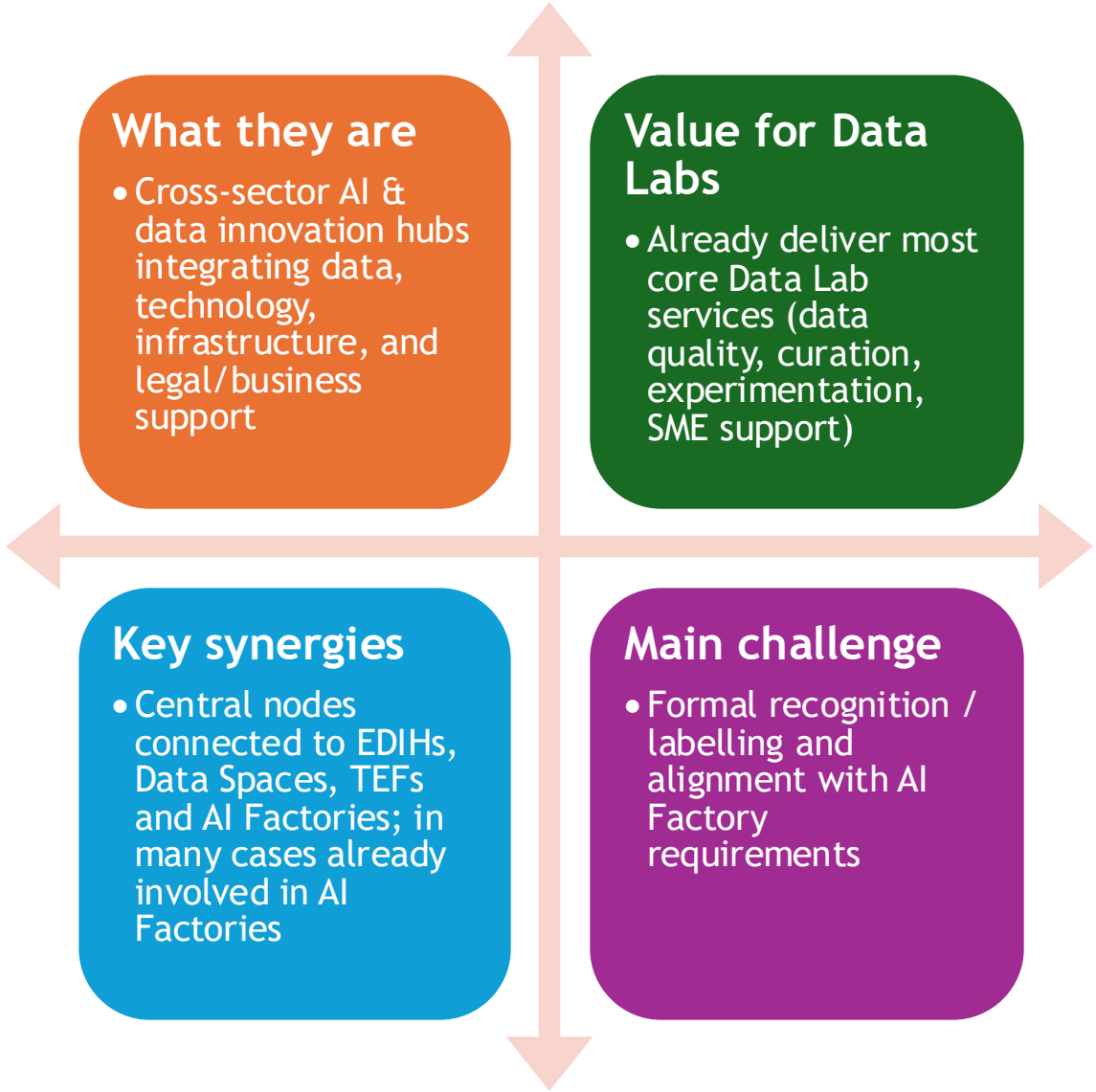
HORIZON-JU-EUROHPC-2026-COAIF-03



Data Labs & BDVA i-Spaces

i-Spaces can fully operate as Data Labs with minimal evolution

Natural Candidates to become Data Labs



Next Steps

Monitoring and evidence building

- Monitor and document early Data Labs implementations, particularly those under AI Factories.
- Collect lessons learned from proofs of concept, pilots, and proof-of-deployment activities

Strategic guidance for future phases

Provide guidance on priorities for future Data Labs, including:

- investment needs,
- opportunities for synergies,
- areas for new pilots,
- needs for reference frameworks and use cases.

BDV BIG DATA VALUE ASSOCIATION

DATA LABS INITIATIVE

DATAWEEK
JOIN.LEARN.SHARE.GET VALUE

ORGANISED BY BDV BIG DATA VALUE ASSOCIATION

SUPPORTED BY The Research Council of Norway

IN COLLABORATION WITH IFE NTNU SINTEF VESTLANDSFORSKING

RI SCALE

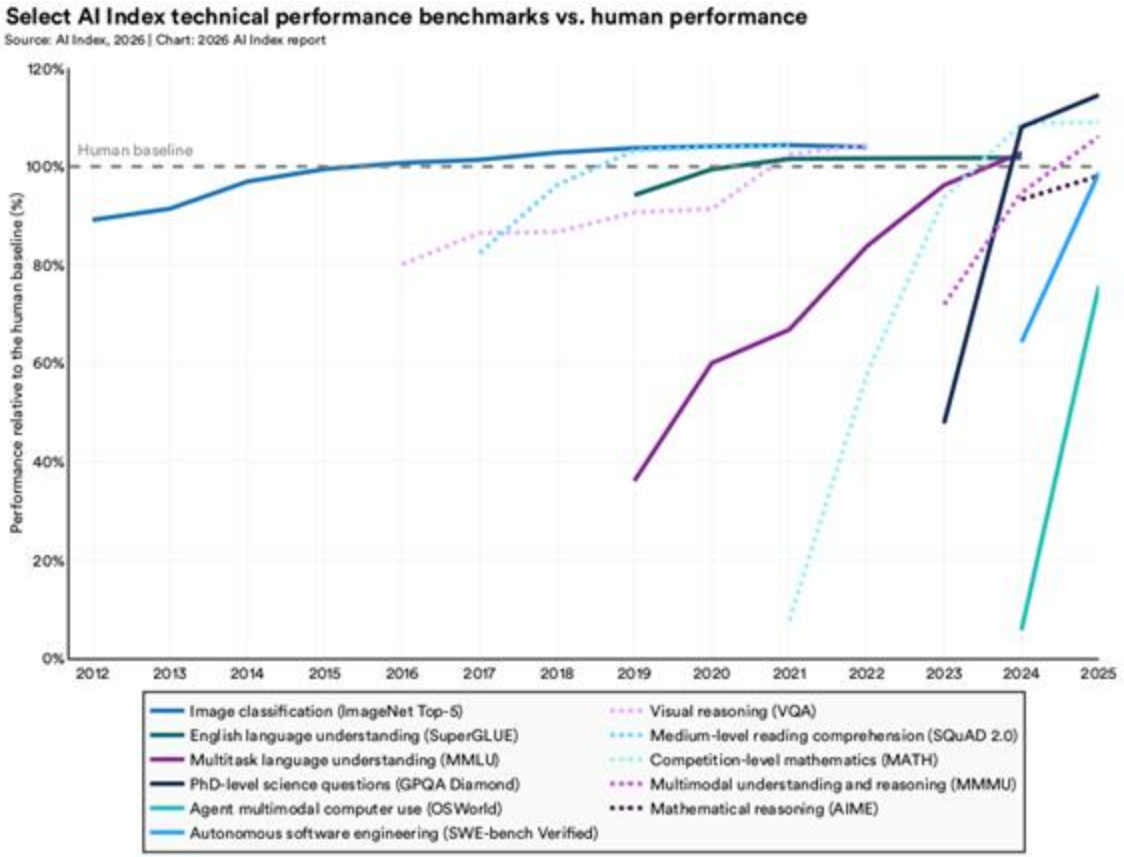
Data Exploitation Platform: tools and opportunities for Data Labs

5.5.2026 Ville Tenhunen, Gergely Sipos
EGI Foundation

Funded by the European Union

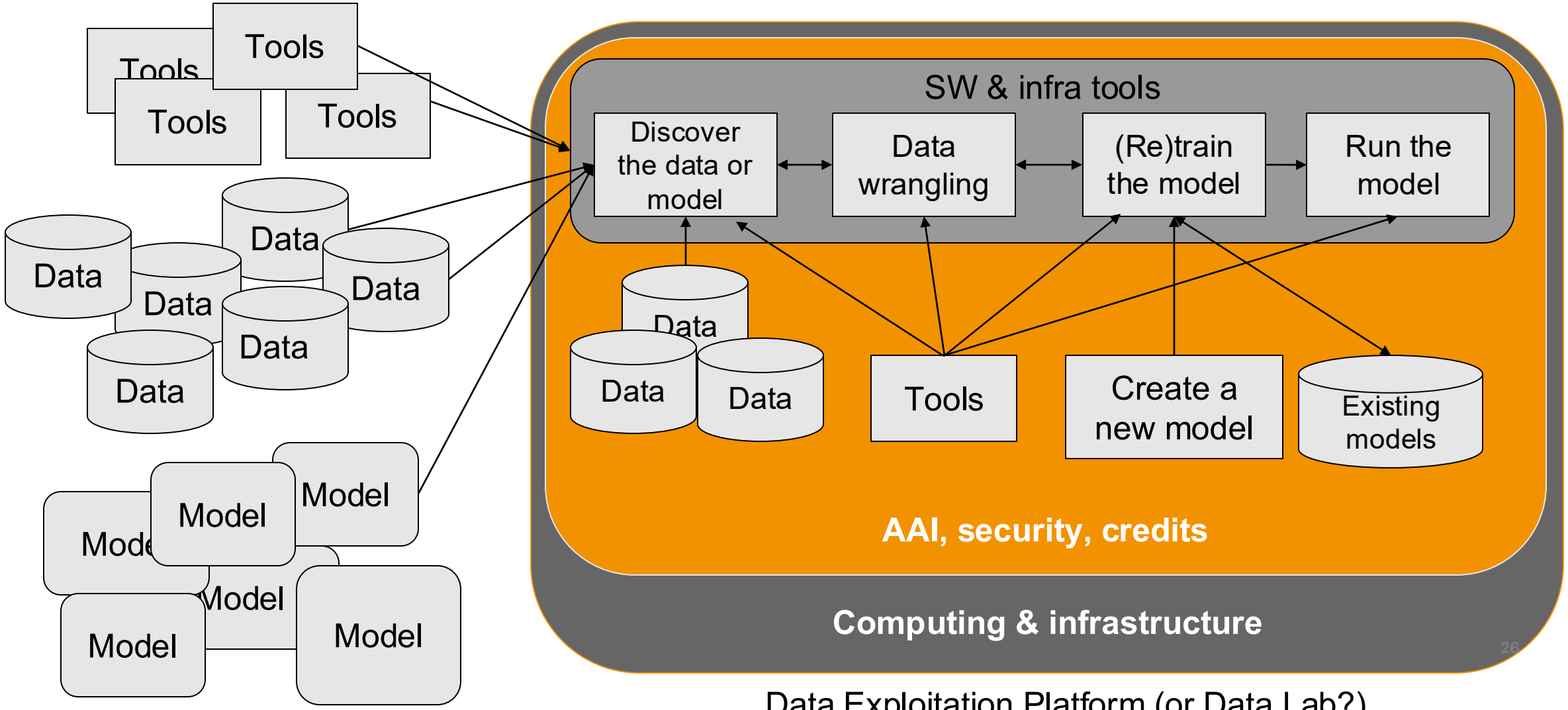
Intro: AI is Shaping Scientific Research

- AI models now meet or exceed human baselines in advanced scientific reasoning, multimodal tasks, and competitive mathematics.
- AI adoption is widespread: 88% of organisations and four out of five university students now use generative AI, with significant impacts on research efficiency and innovation.



Source: Sha Sajadieh, Loredana Fattorini, Raymond Perrault, Yolanda Gil, Vanessa Parli, Lapo Santarlasci, Juan Pava, Nestor Maslej, Russ Altman, Erik Brynjolfsson, Carla Brodley, Jack Clark, Virginia Dignum, Vipin Kumar, James Landay, Terah Lyons, James Manyika, Juan Carlos Niebles, Yoav Shoham, Elham Tabassi, Russell Wald, Toby Walsh, Dan Weld. "The AI Index 2026 Annual Report," AI Index Steering Committee, Institute for Human-Centered AI, Stanford University, Stanford, CA, April 2026.

Workflows Are Not the Same



Data Exploitation Platform (or Data Lab?)

The Challenges of Big Data in Science

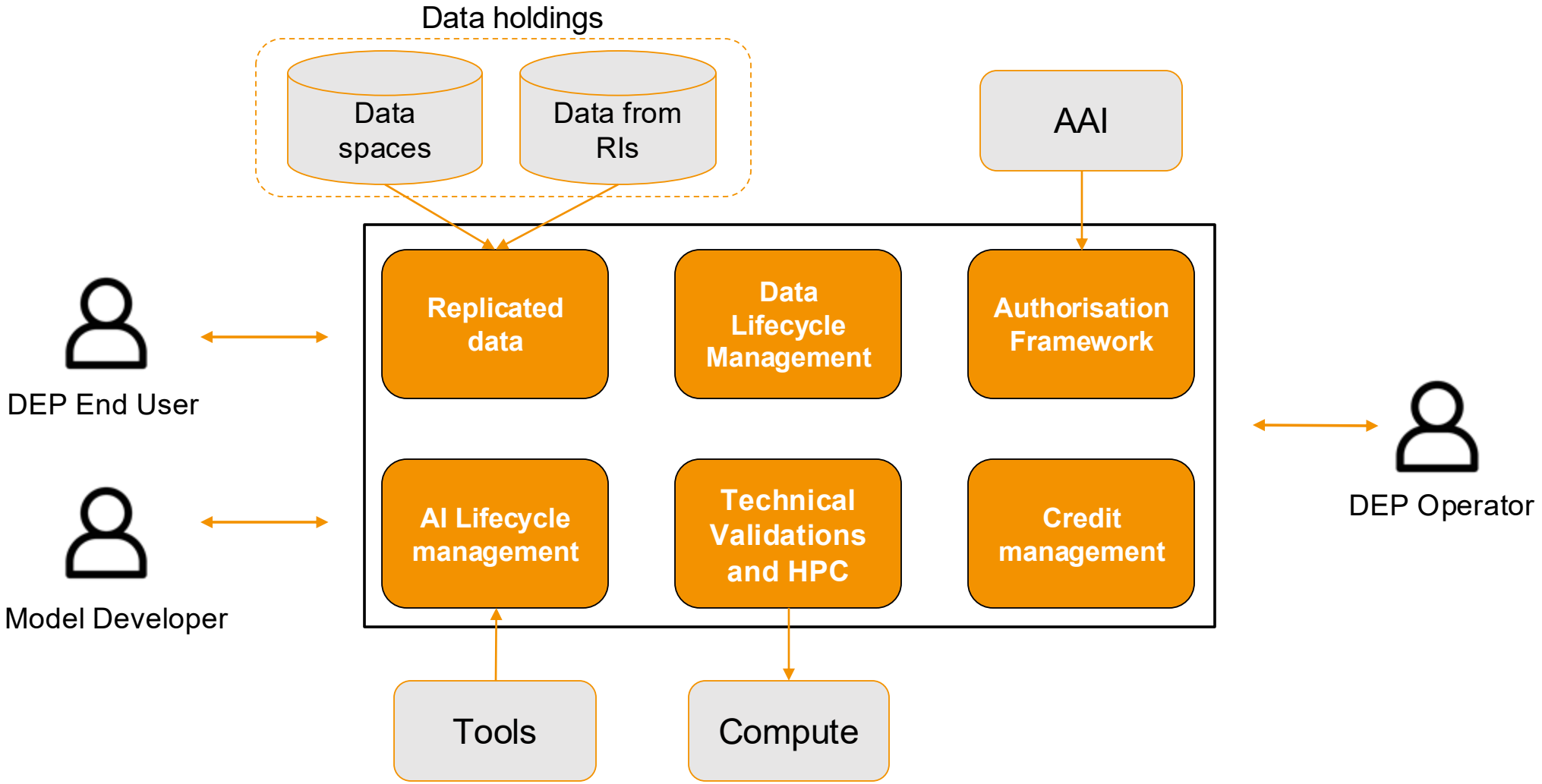
Computing power is where the data isn't

Data access and transfer can be cumbersome

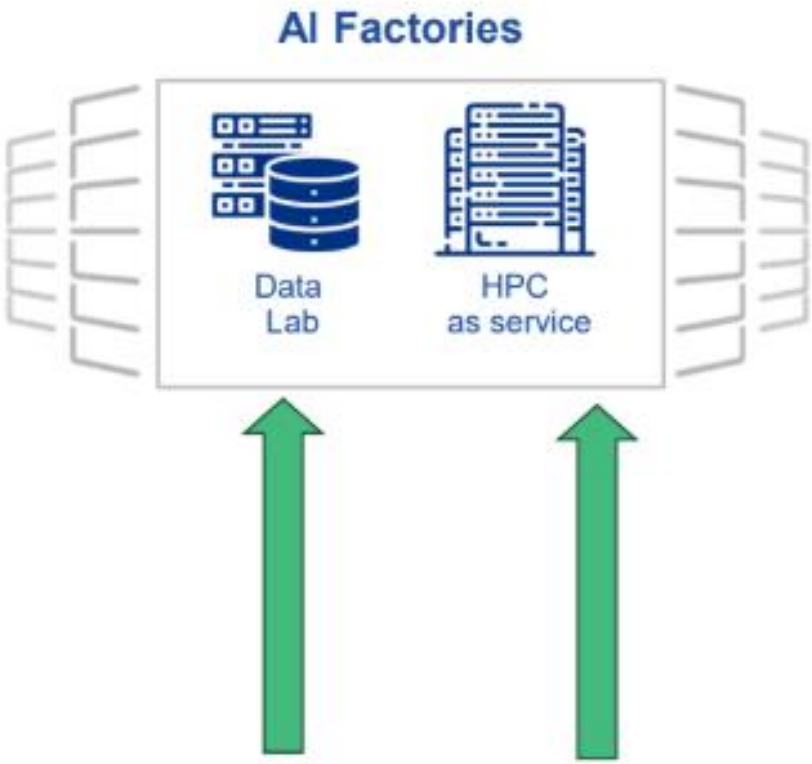
AI tool setup and integration pose high technical barrier

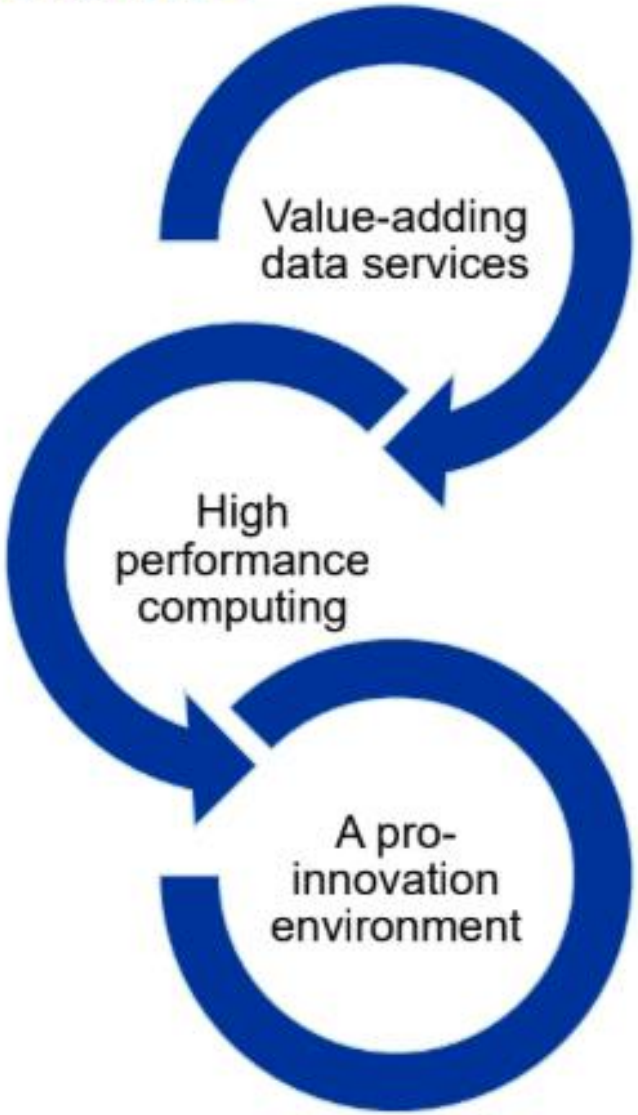
Data Exploitation Platform (DEP)

The backbone for next-gen science data valorisation



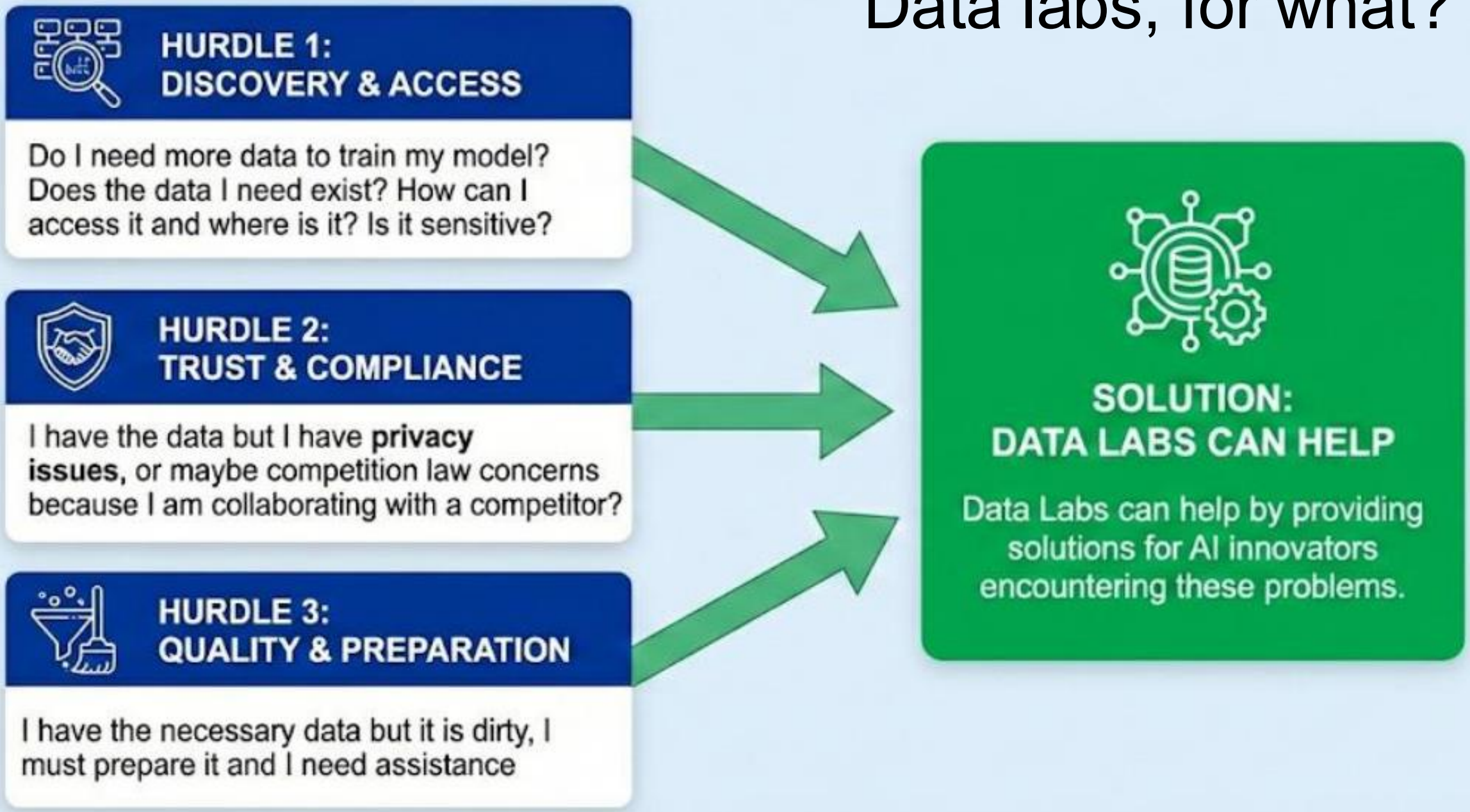
The first instances of Data labs will be deployed at the frontend of AI Factories





Source: Pierre Serra (Policy Officer, EC DG CONNECT, Unit G.1 - Data Policy & Innovation)

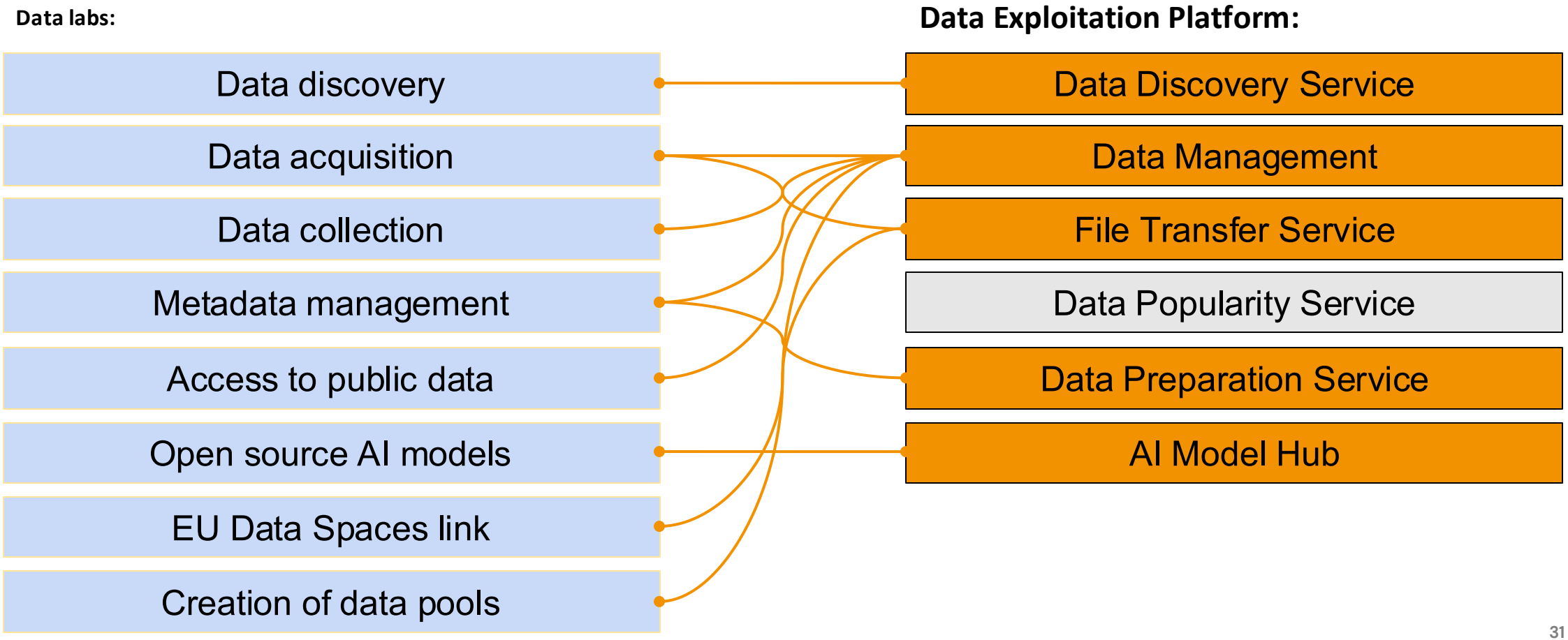
Data labs, for what?



Source: Pierre Serra (Policy Officer, EC DG CONNECT, Unit G.1 - Data Policy & Innovation)

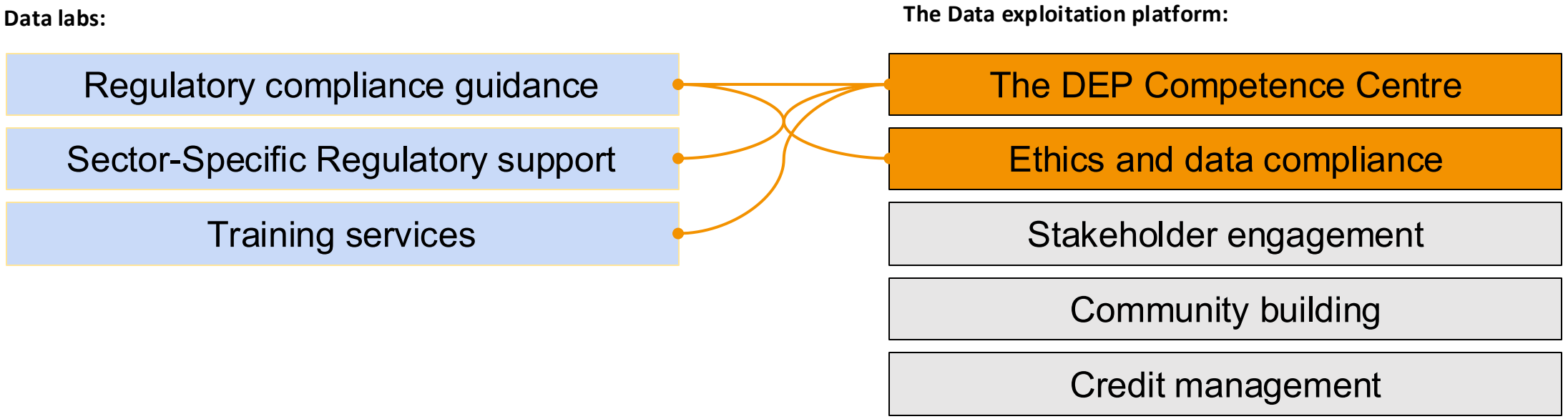
Hurdle 1: Discovery & access

Conceptual mapping



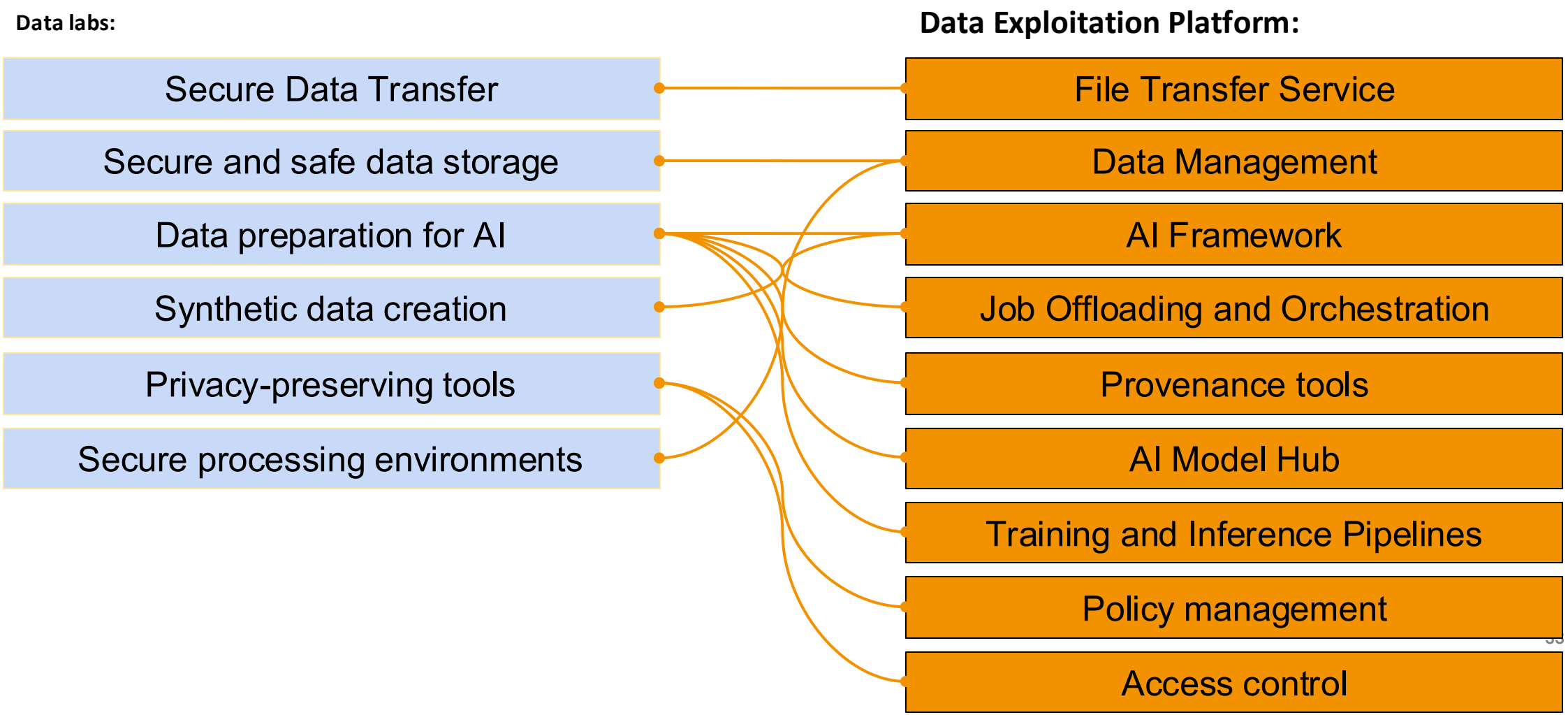
Hurdle 2: Trust and compliance

Conceptual mapping



Hurdle 3: Quality & preparation

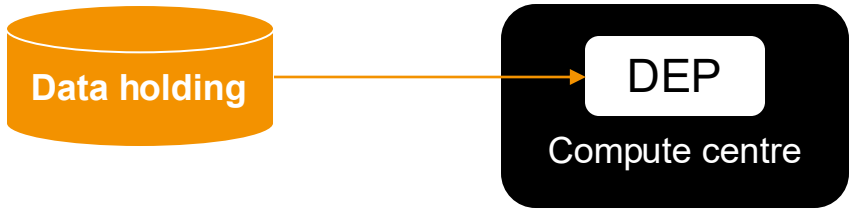
Conceptual mapping



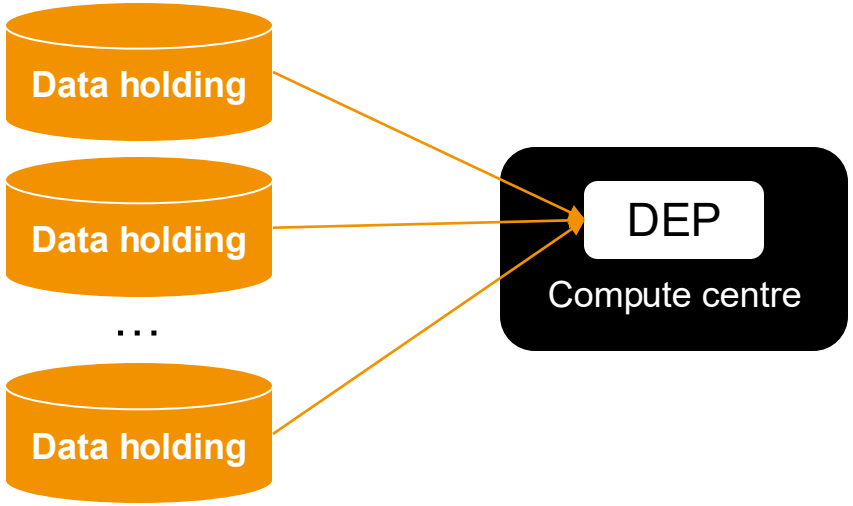
Envisaged operational options

Credit based measurement of usage will help us quantify the cost of operation

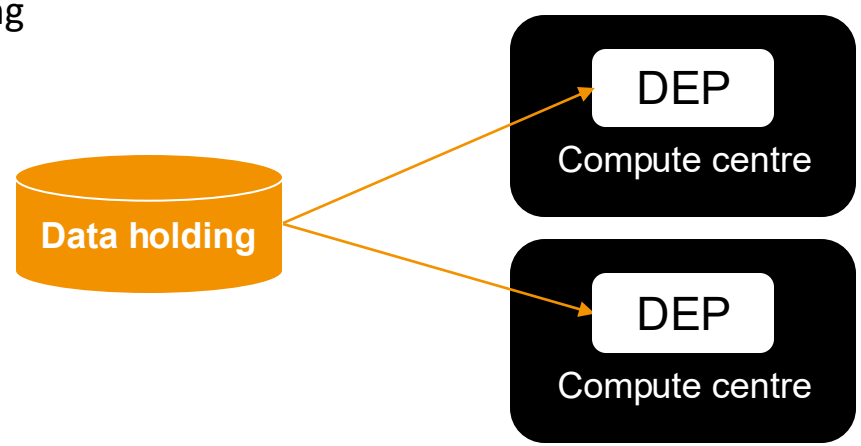
1 DEP is connected to 1 data holding. E.g. ENES



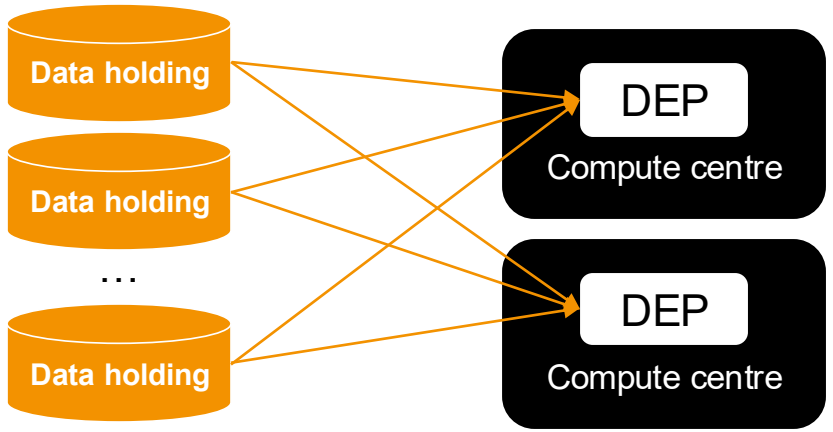
1 DEP is connected to multiple data holdings. E.g. EISCAT



Multiple DEPs are connected to the same data holding. E.g. Euro-BioImaging



Multiple DEPs are connected to multiple data holdings. E.g. BBMRI

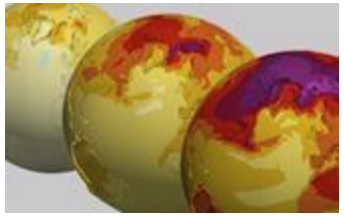


RI SCALE Scientific validation cases for DEP technology

Climate modelling

1. High-resolution **downscaling** of climate scenarios and **risk trend analysis** in agriculture
2. Smart **detection of anomalies** in climate data usage

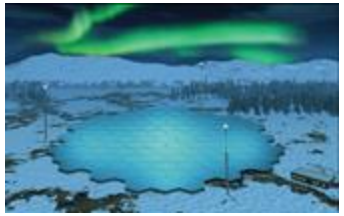
20PB



Near-Earth geospace studies

3. **Intelligent Scheduling** of Radar Observations and Experiments
4. Space **Debris and Anomaly Detection**

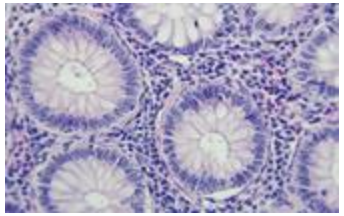
3PB



Cancer treatment

5. Colorectal **cancer prediction** with explainable AI
6. **Synthetic data** for computational pathology

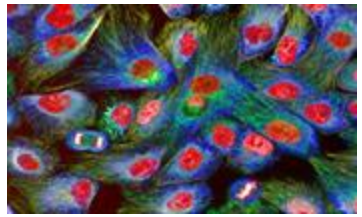
3PB



Imaging in life sciences

7. **Foundational models** for heterogeneous biological image data
8. **Generative AI-Powered assistant** for data discovery and analysis

5PB



RI SCALE Technological validation cases for the DEP

Destination Earth ECMWF

Scalability on EuroHPC

Our models can run on Europe's largest supercomputers for Destination Earth, the highly accurate 'Digital Twin of the Earth'

BBMRI Biobanking and Biomolecular Resources Research Infrastructure

Masaryk Memorial Cancer Institute

Image compression

Achieve DEP performance improvement by using JPEG2000 Part 2 compression in DICOM image-metadata format

fragmentiX QUANTUM AGE CYBERSECURITY

TU WIEN TECHNISCHE UNIVERSITÄT WIEN Vienna | Austria

Masaryk Memorial Cancer Institute

Green computing

Achieve energy saving and performance improvement with the use of GROQ cards for AI model training and inference

egi

grnet

TÜBİTAK

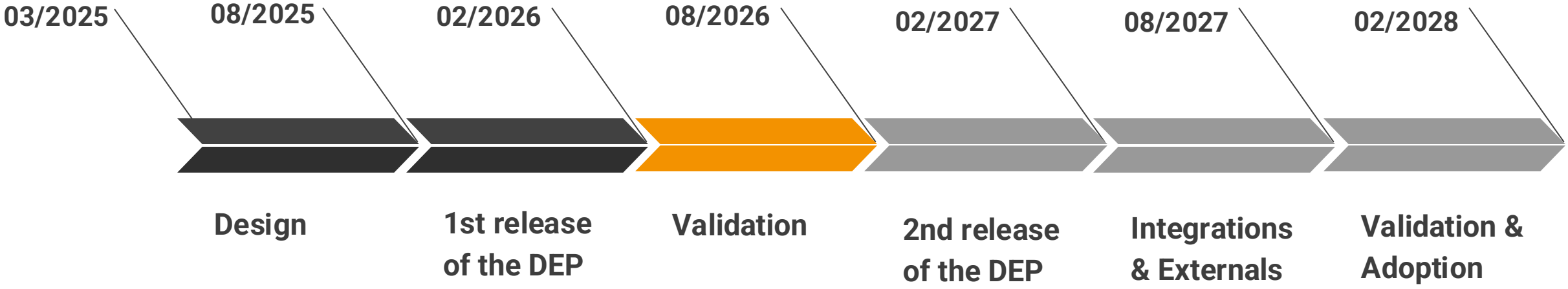
TU WIEN TECHNISCHE UNIVERSITÄT WIEN Vienna | Austria

Credit management

Be able to measure Data Exploitation Platform usage and environmental impact, and express them in form of 'credits'

RI-SCALE in a Nutshell

- Call: HORIZON-INFRA-2024-TECH-01-01
- Coordinator: EGI Foundation
- 29 partners: Research Infrastructures, e-Infrastructures, AI and Data Developers
- Budget: €10,5M (funded by EC)
- Duration: From 1.3.2025 to 29.2.2028



Conclusions

- The Data Exploitation Platform **concept covers** most Data Lab functionalities
- RI-SCALE has **defined requirements and use cases**, including 12 existing scientific use cases for the concept
- The project also analyses and identifies **sustainable business models** and develops **operational model** for the DEP

Get involved
www.riscale.eu/get-involved

Thank you!
Let’ talk.

https://riscale.eu/

Funded by
the European Union

RISCALE

DATAWEEK

JOIN.LEARN.SHARE.GET VALUE

Data Spaces – Data Labs – AI Factories
A match made in heaven!

Jürgen Vangeyte (ILVO, CEADS data space)

ORGANISED BY BDV BIG DATA VALUE ASSOCIATION

SUPPORTED BY The Research Council of Norway

IN COLLABORATION WITH IFE NTNU SINTEF WESTERN NORWAY RESEARCH INSTITUTE VESTLANDSFORSKING

CEADS COMMON VISION

Layer 1: CEADS collaboration framework

Layer 2: CEADS services

Authority services

Non-Authority services

Offered by CEADS

Offered by DSI or service providers

Research and innovation data for improved AI algorithms performance

Layer 3: Data and Application services

UC based services

- **Semantic enrichment => Making data discoverable and (re)usable**
  - Ontologies (AGROVOC)
  - Standard metadata (DCAT-AP)
- **Data annotation => Making data AI-ready (pipeline) A**
  - Annotating raw data (e.g. drone images → masks → semantic labels)
  - Harmonizing labels rather than inventing new ones (important governance point)
  - Structured datasets

  => AI-ready pipelines
- **Data harmonization=> Making data interoperable across across DSIs**
  - Cross-platform interoperability
  - Harmonizing datasets using shared models (e.g. Agricultural Information Model) –
  - => Enable cross-data-space queries and services

DATA LAB SERVICES

From UC based to reusable, scalable services

Robotic data

Satellite data

Sustainability data

Food manufacturing data

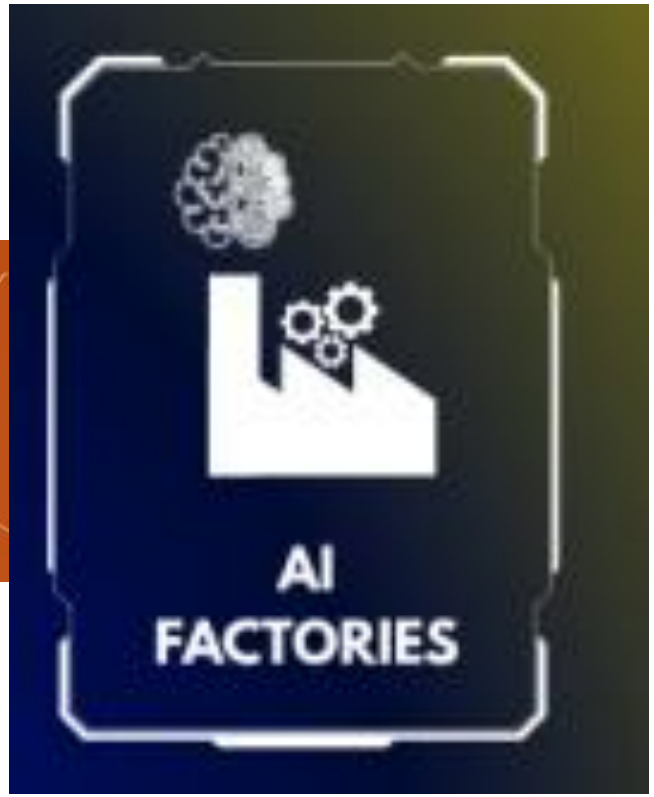
Machinery data

Testing data sets (TEFs)

.....

DATA LAB SERVICES

DATA LAB SERVICES



MATCH?

Data Labs services help unlock the value of Data spaces

DATAWEEK
JOIN.LEARN.SHARE.GET VALUE

AI REDGIO 5.0

Data Labs: their role in Europe's Data and AI Landscape

the EU Digital Innovation Hubs viewpoint

Sergio Gusmeroli (MADE Competence Center)

ORGANISED BY BDV BIG DATA VALUE ASSOCIATION
SUPPORTED BY The Research Council of Norway
IN COLLABORATION WITH IFE NTNU SINTEF VESTLANDSFORSKING

The European Network of DIH (EDIH)

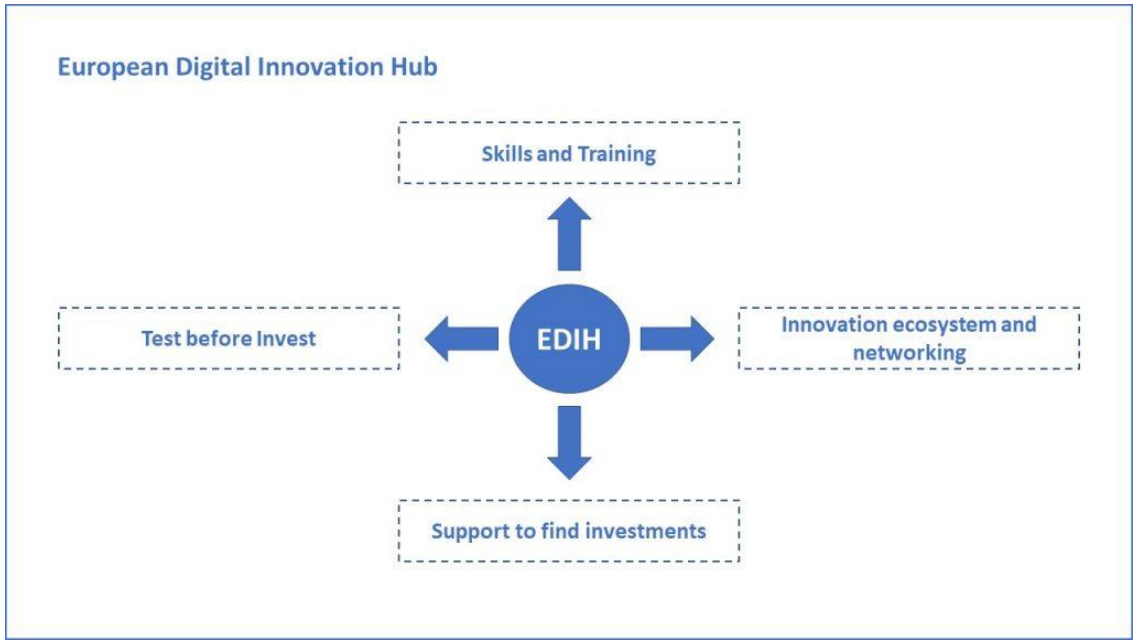
Webtools + © EC-GISCO + Leaflet | © OpenStreetMap contributors | Avvertenza

Funded under Digital Europe Programme

Seal of Excellence

EDIH
European Digital Innovation Hubs Network

255 EDIHs



The EDIH2.0 in the EU AI Ecosystem

AI REDGIO 5.0

AI-on-Demand Platform
is a central (online) marketplace for AI resources.

AI Skills Academy
attracts talent, upskills and reskills students and professionals in the EU.

AI REDGIO 5.0

Provide tools and AI resources to

EDIHs
are one major point of entry and information for users and providers of AI systems.
The AI Service Desk, collaborating with the network of EDIHs, will answer AI Act related questions.

EDIH4Manu 2.0

Redirect stakeholders to support infrastructure

ITALIA AI FACTORY

AI Factories
provide necessary resources (computing, skills and services) to develop advanced AI models.

TEFs
help innovative tech providers bring products from lab to market.

may use technical support and infrastructure of

AI regulatory sandboxes
provide a special regime and legal guidance for testing in real-world conditions.

EUSAiR
BRIDGING INNOVATION AND COMPLIANCE IN AI

AI REDGIO 5.0

AI matters

TEF agrifood

The D BEST taxonomy (Data driven DIHs)

D — Data

- DATA ACQUISITION & SENSING
- DATA PROCESSING & ANALYSIS
- DECISION MAKING
- PHYSICAL–HUMAN ACTION & INTERACTION
- DATA SHARING

B — Business

- INCUBATION ACCELERATION SUPPORT
- ACCESS TO FINANCE
- OFFERING HOUSING
- BUSINESS TRAINING AND EDUCATION
- PROJECT DEVELOPMENT

E — Ecosystem

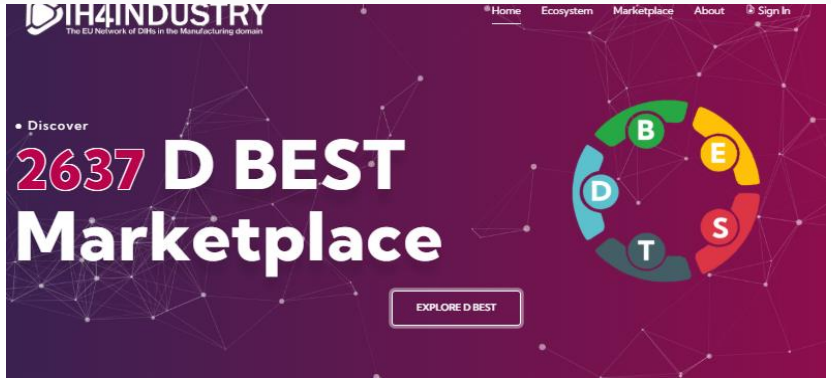
- COMMUNITY BUILDING
- DIH INNOVATION DEVELOPMENT
- ECOSYSTEM GOVERNANCE

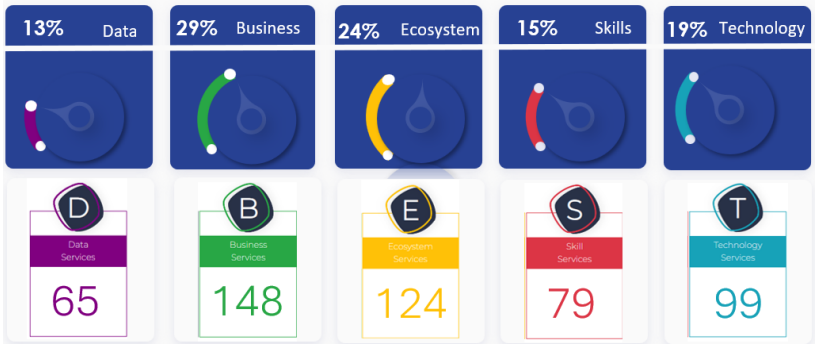
S — Skills

- PROCESS & ORGANIZATIONAL MATURITY
- HUMAN CAPABILITY MATURITY
- SKILLS IMPROVEMENT

T — Technology

- IDEAS MANAGEMENT & MATERIALIZATION
- CONTRACT RESEARCH
- PROVISION OF INFRASTRUCTURE
- TECHNICAL SUPPORT ON SCALE UP
- VERIFICATION AND VALIDATION





CLASS	TYPE	SERVICE
DATA	Data acquisition and sensing	Data acquisition
		Data protection
	Data processing & analysis	Data storage
		Data analytics
	Decision-making	Cognitive Big Data architectures
		Decision support and development
	Physical-human action & interaction	Collaborative intelligence
		User experience
		Feedback loop
	Data sharing	GDPR and Data Sovereignity compliance
		Data Spaces
		Data Platform

Five good reasons for EDIHs to become Data Labs

1. **Long-lasting** multi-disciplinary and neutral **entities** (since **2016** DEI communication) in EU regional innovation ecosystems
2. **Capillary** decentralised **presence** in more than **255 EU Regions**
3. **Recognised & Trusted Authority** for the associated members (no-profit, multi-vendor, open, interoperable) especially SMEs
4. Recent extension (AI REDGIO5.0) of their original scope (BEST) to Data-related services (**D BEST**) and close collaboration with **Data Spaces**
5. **EDIH 2.0** more technical focus in the EU **AI Ecosystem** of AI Factories, AI TEFs, AI Skills Academy, Regulatory Sandboxes & AIonDemand Platform

DATAWEEK
JOIN.LEARN.SHARE.GET VALUE

AI REDGIO 5.0

THANKS!!!

The EU Digital Innovation Hubs viewpoint

Sergio Gusmeroli (MADE Competence Center)
sergio.gusmeroli@made-cc.eu

ORGANISED BY BDV BIG DATA VALUE ASSOCIATION
SUPPORTED BY The Research Council of Norway
IN COLLABORATION WITH IFE NTNU SINTEF VESTLANDSFORSKING