DATAWEEK

JOIN.LEARN.SHARE.GET VALUE

EUH4D - EuroCC Synergies and alignment (bringing HPC-Data-AI ecosystems closer together)

EUHUBS4DATA

BDV BIG DATA VALUE ASSOCIATION

Daniel Opalka, EuroHPC JU, EC

Introduction to how EuroHPC JU expects the national competence centres to be able to meet the goals set for Europe

EuroHPC
Joint Undertaking

EUROCC – A pan-European network of National Competence Centres for high performance computing

Daniel Opalka
Head of Research & Innovation

Support SMEs and industry with HPC

EuroHPC Joint Undertaking

Expertise and HPC resources

CASTIEL

EURO

- European network of NCCs in 32 countries to **widen the use of HPC** in Europe.
- Support **SMEs, public services** and **private businesses**
- **Gateway for** users to access **European HPC** resources

Financial support for HPC uptake

- Boost **innovation** and **business opportunities for SMEs**
- **Solve business problems with HPC,** financial support and expertise from European HPC leaders

f4 EuroHPC

Talent development

EUMaster4HPC

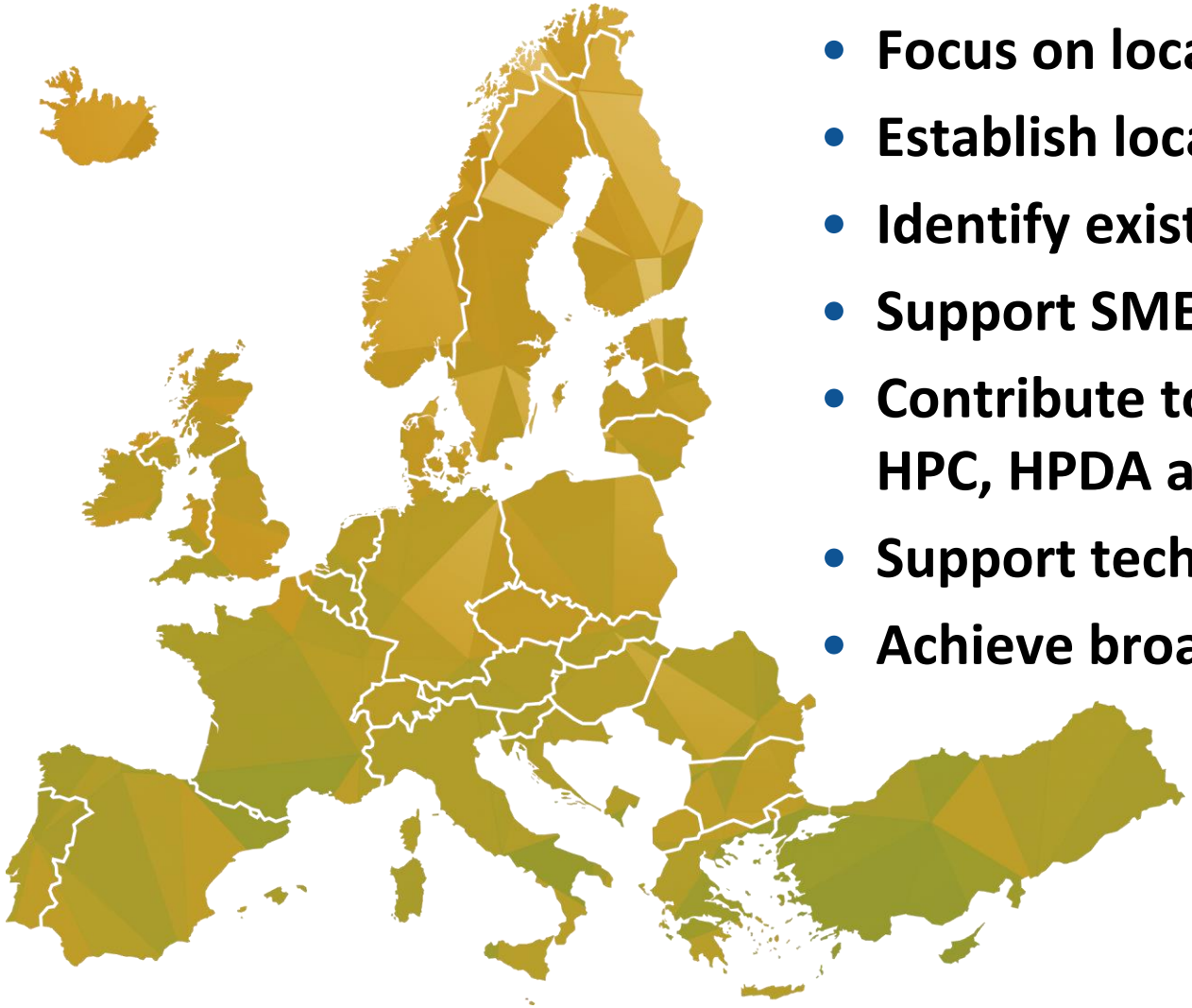
- Train the **next generation of HPC experts** in Europe
- 1st **pan-European MSc** Programme in HPC including 19 countries
- **Connect academic education with HPC industry**

„ It is our proposed level of ambition that by 2030:

- *75% of European enterprises have taken up cloud computing services, big data and Artificial Intelligence*
- *More than 90% of European SMEs reach at least a basic level of digital intensity “*

EUROCC - Mission and key objectives

A pan-European network of National Competence Centres in HPC



- Focus on local/national activities and communities
- Establish local/national focal points of HPC expertise
- Identify existing local and national competences
- Support SMEs with knowledge, expertise, training
- Contribute to tailored national training activities in HPC, HPDA and other HPC-related fields
- Support technology transfer at local/national level
- Achieve broad geographical coverage

National Competence Centres for HPC

- **EUROCC / EUROCC 2**
  A European network of more than 30 national HPC Competence Centres to widen the use of HPC in Europe
- **CASTIEL / CASTIEL 2**
  Coordinate and support the NCC network (and the European Centres of Excellence for HPC Applications in phase 2), e. g. with twinning and mentoring initiatives

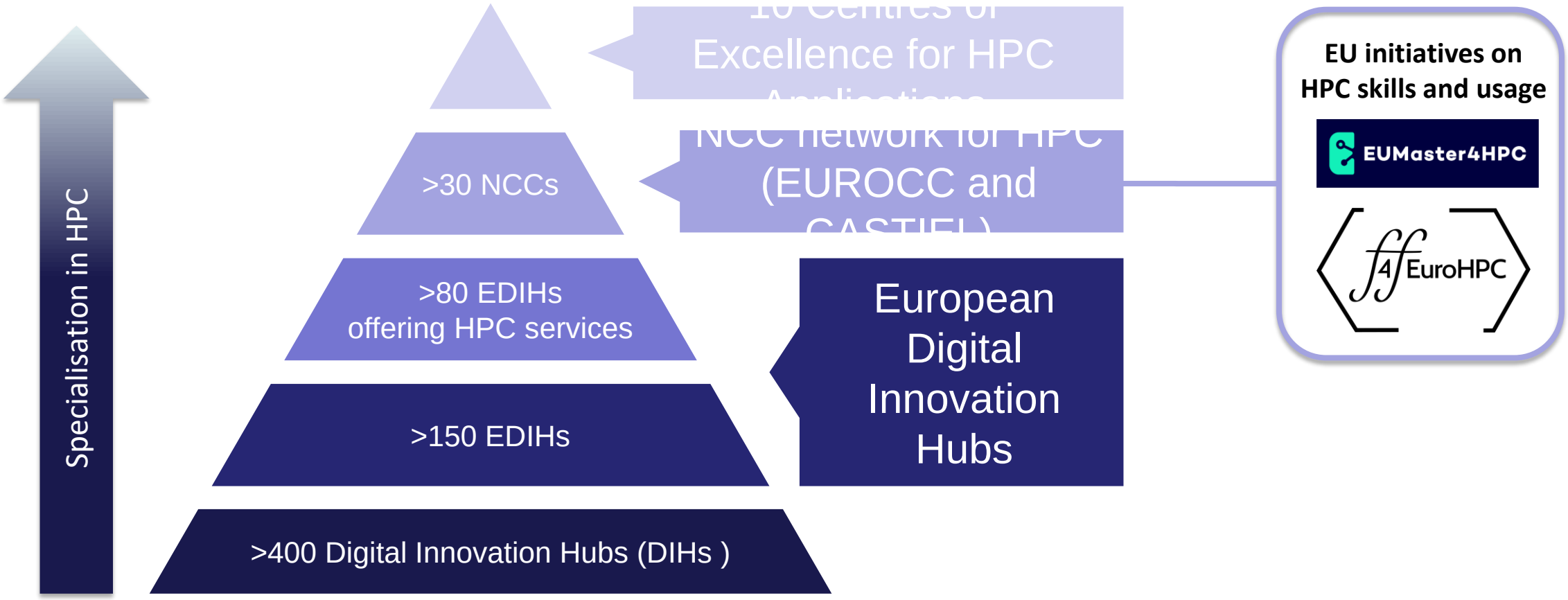
https://www.eurocc-access.eu/

EUROCC / CASTIEL
09/2020 – 12/2022

EUROCC 2 / CASTIEL 2
01/2023 – 12/2025

EUROCC in the European HPC context

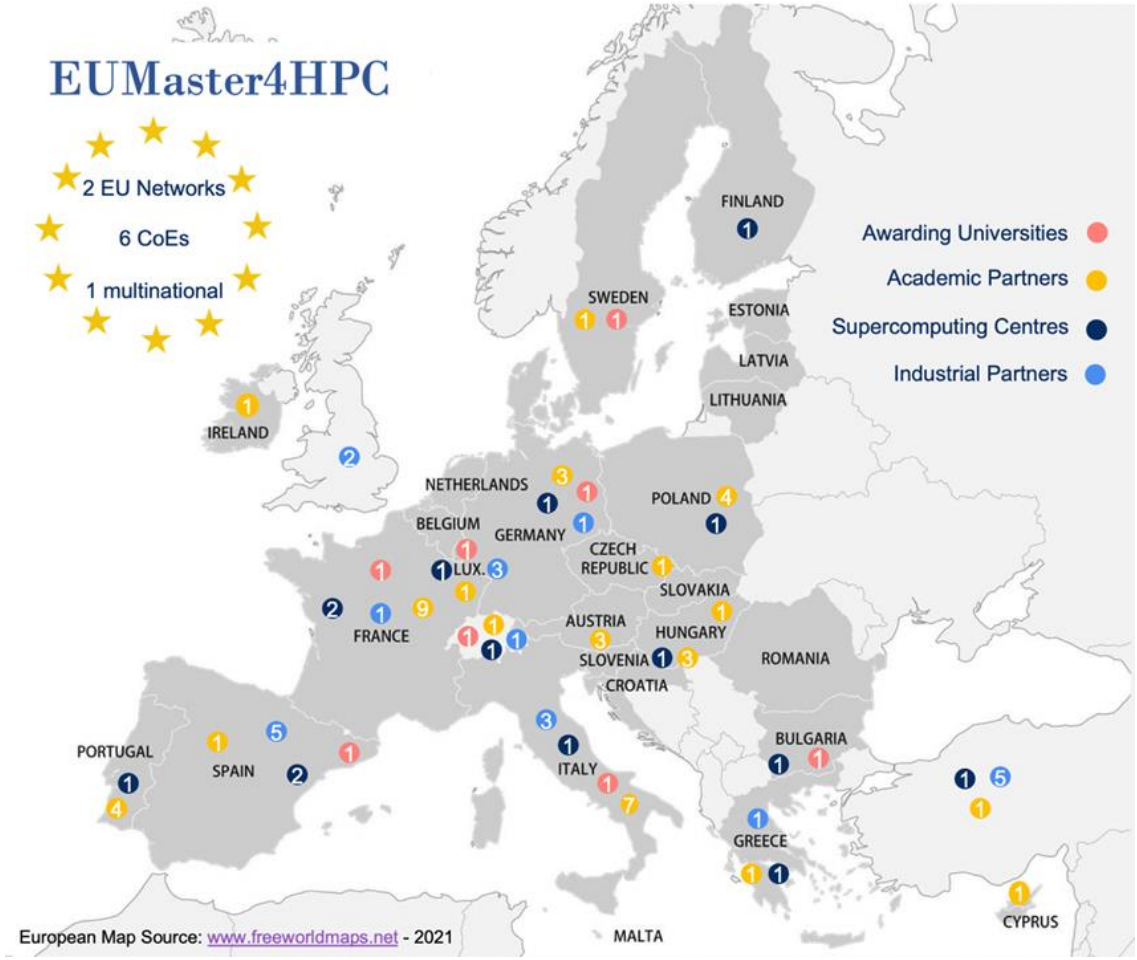
Promoting HPC uptake beyond basic digitalisation of businesses



EUMaster4HPC – a European MSc programme

Educating the next generation of HPC professionals

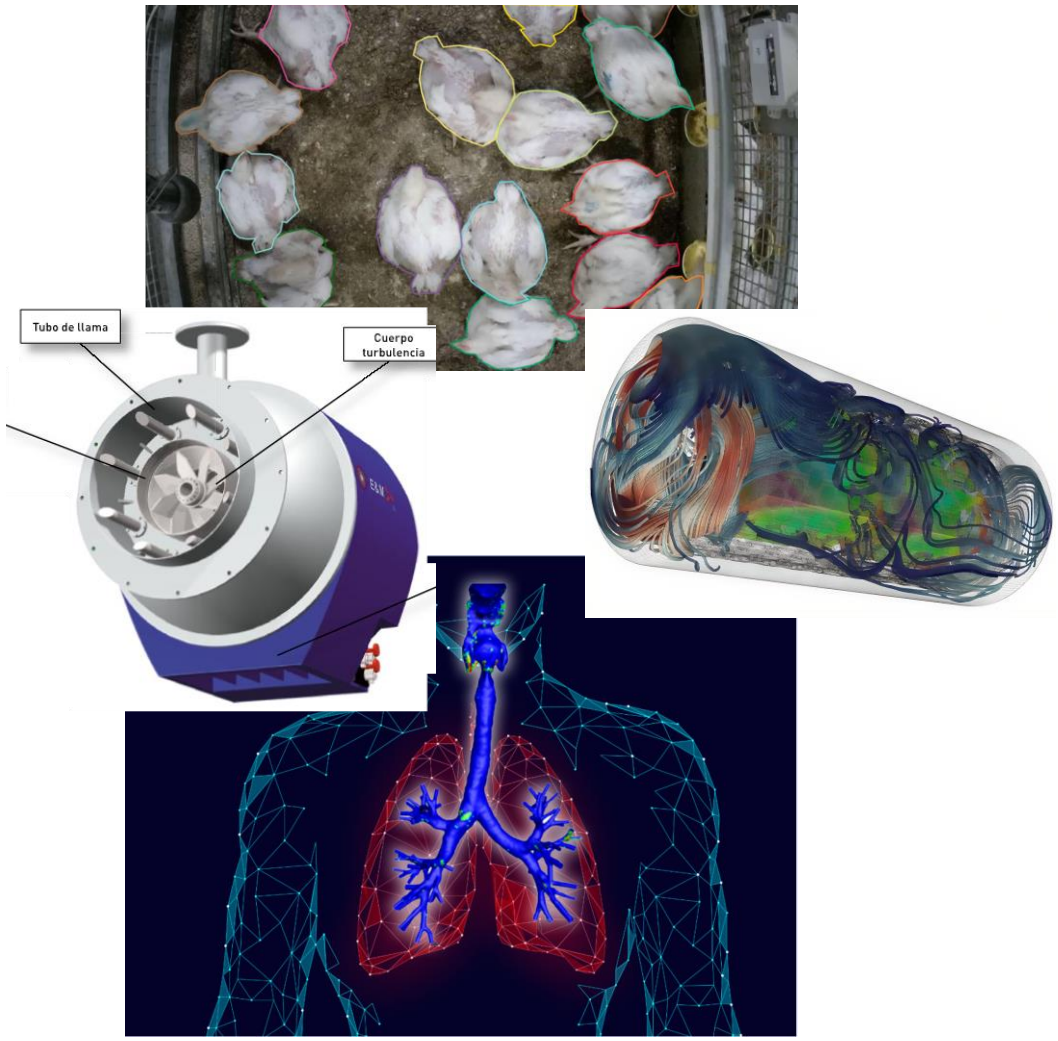
- **European MSc Programme for HPC**
  - **Design of a common European master reference curriculum in HPC and its implementation in a pilot education programme**
  - **Catalyse HPC adoption by educating the next generation of HPC talents for industry, academia and entrepreneurship**
- **Create an open & sustainable network of stakeholders**
  - **Higher education institutions**
  - **Public organisations with interest in HPC**
  - **Private companies operating in an HPC context**



https://eumaster4hpc.uni.lu/

Stimulating the innovation potential of SMEs

https://www.ff4eurohpc.eu/



- **Financial support for SMEs to**
  - **Boost innovation potential and competitiveness through HPC**
  - **Address business challenges with HPC**
  - **Adopt HPC in existing business models**
  - **Develop skills in related disciplines on the basis of HPC, such as high performance data analytics (HPDA) and artificial intelligence to solve specific business problems**
- **Grants of 50.000 – 150.000 € per beneficiary**
- **Support by experts from leading European HPC centres**

DANIEL SÁEZ DOMINGO

Presentation of EUHubs4Data project and i-Spaces

DATAWEEK

JOIN.LEARN.SHARE.GET VALUE

i-Spaces and EUHubs4Data project

Daniel Sáez-Domingo

Technology Transfer Director in ITI
Coordinator EUHubs4Data
BoD member in BDVA

ITI INVESTIGATE TO INNOVATE

EUHUBS4DATA

BDV BIG DATA VALUE ASSOCIATION

BDVA i-Spaces:
https://www.bdva.eu/i-spaces-2022

BDVA i-Spaces 2022

37 hubs
35 regions
21 countries

EUHUBS4DATA

EUROPEAN FEDERATION OF DATA DRIVEN INNOVATION HUBS

PROBLEM?

Data-Driven Innotation Hubs

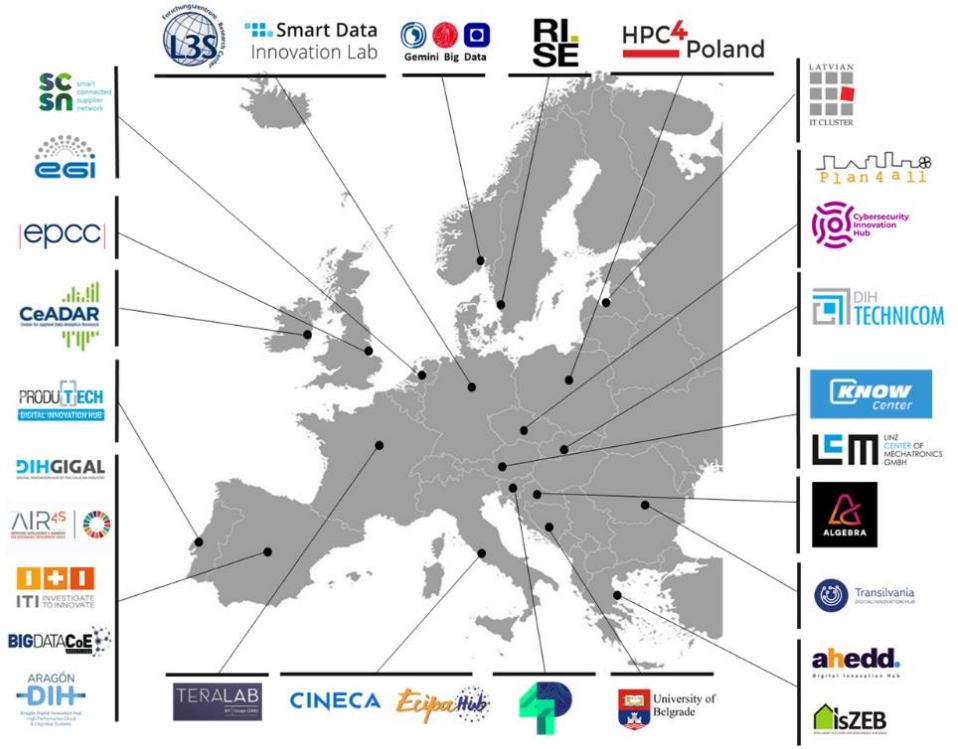
Lack of collaboration to fully exploit their capacities

WHAT'S OUR VALUE?

The major EU initiative focused on Data-Driven experimentation, based on the cross-border collaboration among the hubs.

A pan-european federated catalogue (datasets, services and courses) for data experimentation

Very practical experiences from a legal, ethical and data management perspective on 42 experiments with SMEs and innovation hubs.



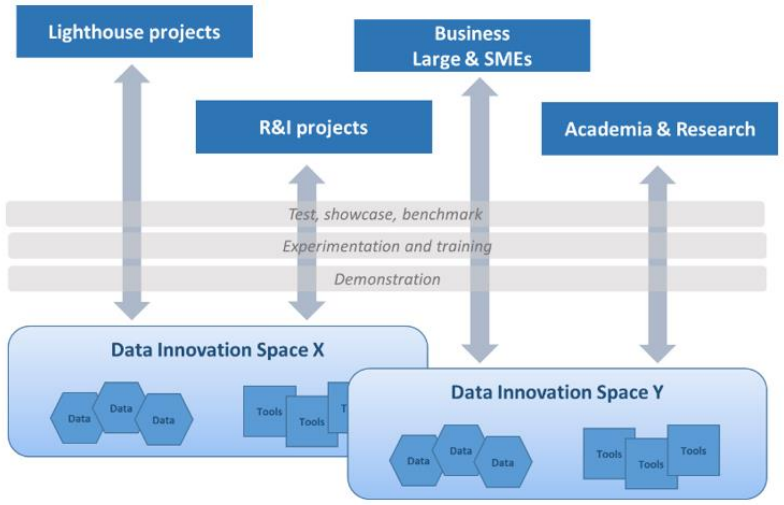
37
i-Spaces / hubs

21
Countries

Federation of i-Spaces/hubs working collaboratively and contributing to the creation of European common Data Spaces.

Coordinator: ITI
Budget: 12.5M
40 Months (sept 2020 – dec 2023)

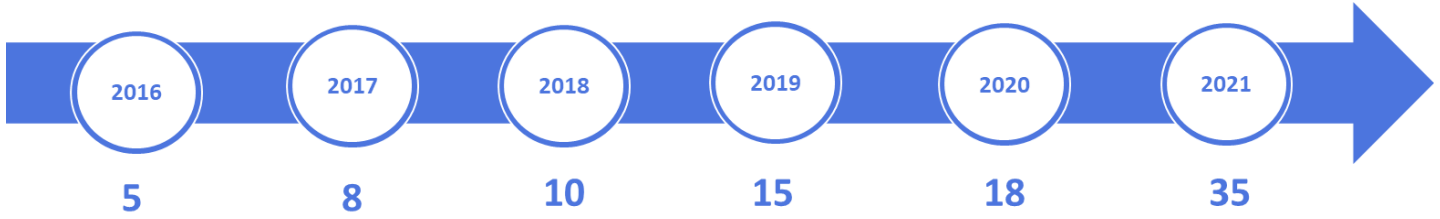
BDVA i-Spaces, the core of the Federation



Federation is in the best position to give **QUALITY** to experimentation, innovation and deployment of Data Technologies

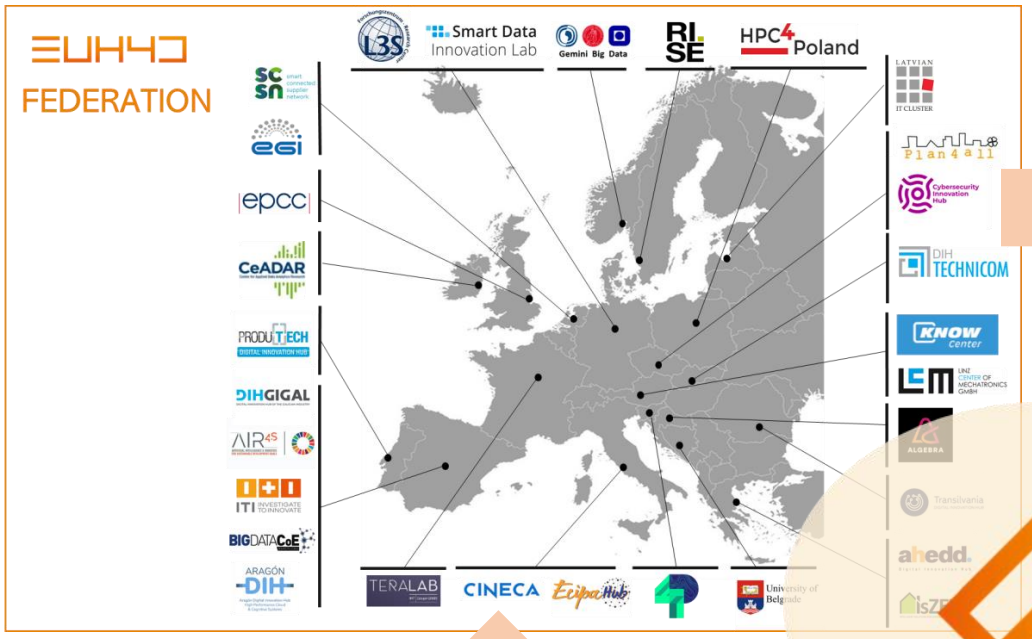
- Long history of collaboration inside BDV PPP
- Shared best practices
- Complete and balanced catalogue
- European coverage
- Joint infrastructures, services, people
- Reference environment for data driven testing and experimentation
- Sustainable structure

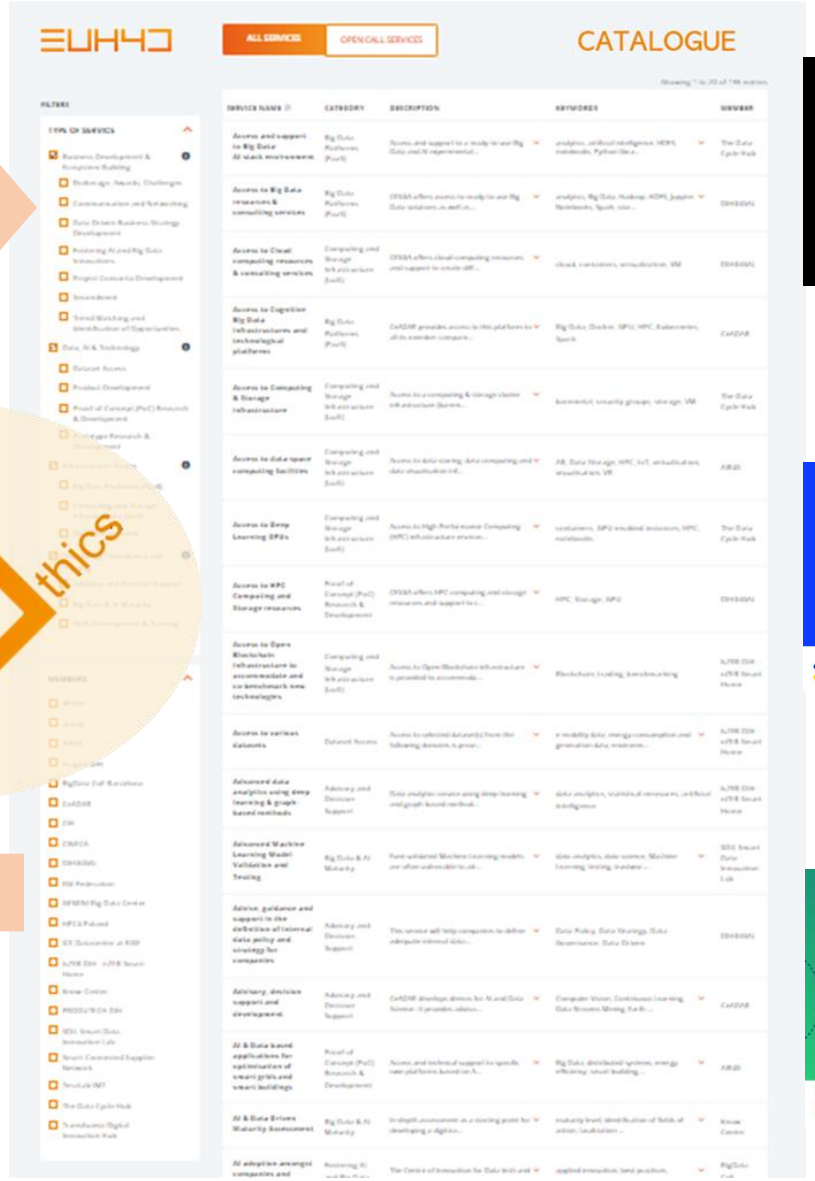
EUHUB
EUROPEAN FEDERATION OF DATA DRIVEN INNOVATION HUBS

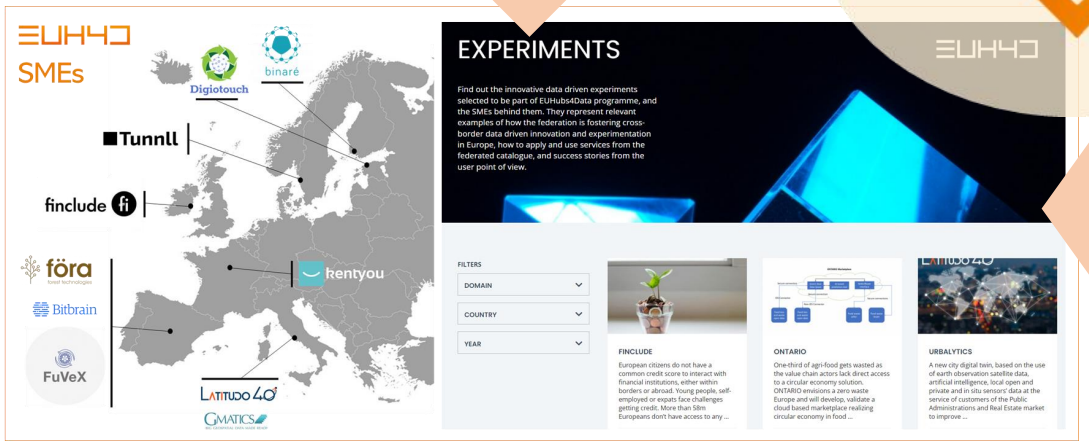


Federation of i-Spaces / data driven hubs

5 main pillars



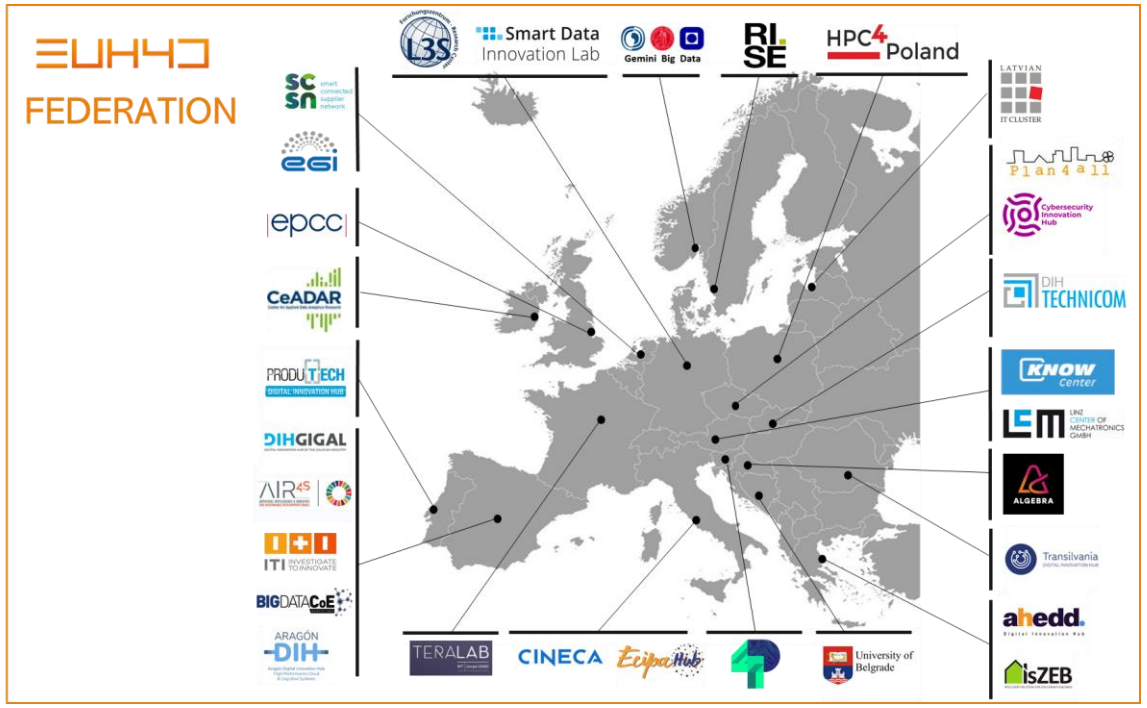




Ethics

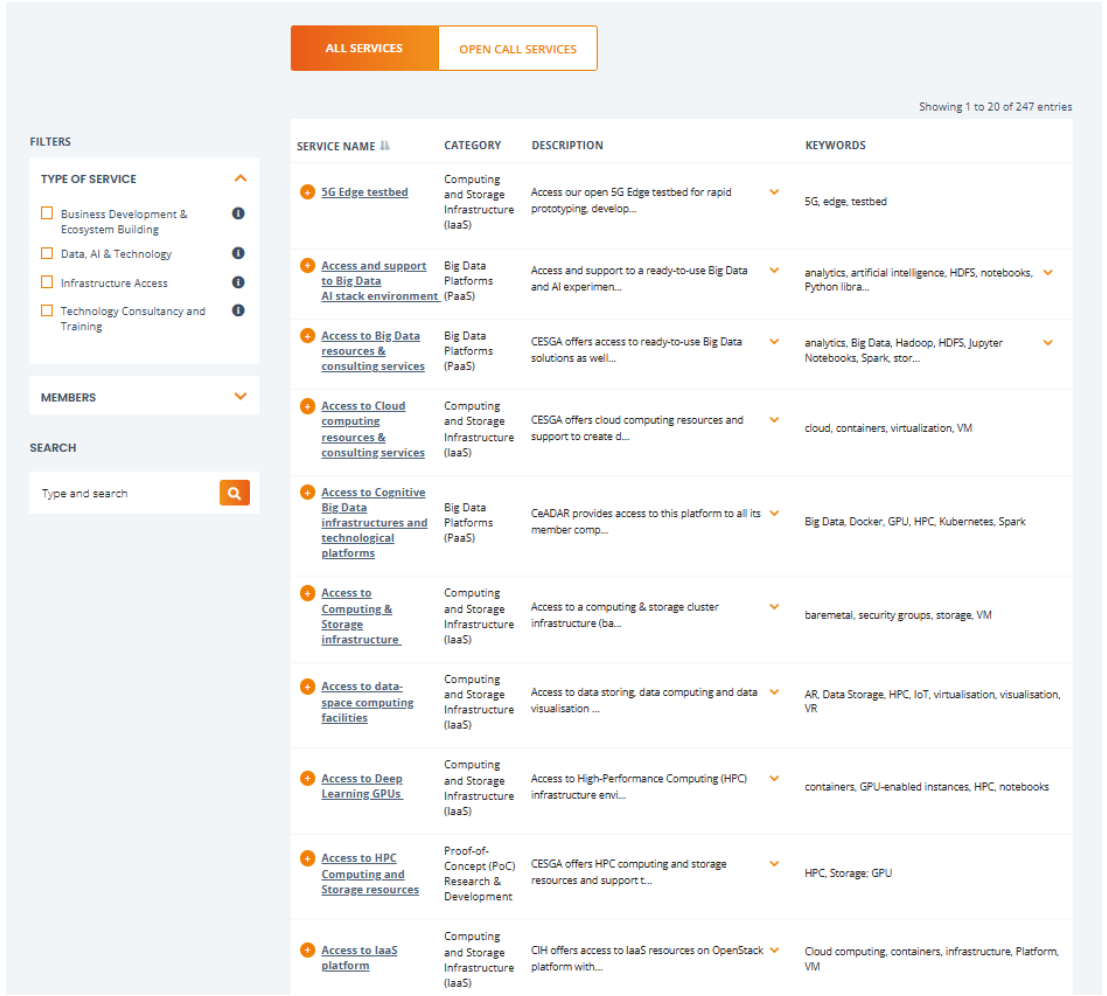
The federation

- 12 DIHs partners covering 12 regions in 10 countries
- 18 new DIHs covering 18 new regions and 10 new countries
- Basis for collaboration among all DIHs already established



- New DIHs involved in catalogue and experiments
- Federation Working Group to drive the establishment of the federation created
- First version of governance and organizational model

The Offer: Services in EUHubs4Data



SERVICES

ITI, DATA GENERATOR

TYPE	REGION	LANGUAGE
Data, AI & Technology, Proof-of-Concept (PoC) Research & Development	Spain	English, Spanish

Access to a data generator that allows you to automate the creation of a dataset based on criteria provided by the user

SERVICE DESCRIPTION

Access to a distributed synthetic data generator, based on spark, capable of generating large amounts of data in very optimal times.
It allows the creation of a set of information in an automated way, based on certain criteria provided by the user, among which are; types of data, data characteristics, randomness, temporality, etc.
Consumable through API in a very simple way, obtaining the results in a persistent volume with capacity to download locally the result ready for its exploitation.

SPECIAL ACCESS CONDITIONS

Access is provided by API

PREREQUISITES

The user must be registered

CASE EXAMPLES

A user wants to perform data analysis in the context of AI and the necessary information is not available for it. With the data generator the user can create this data synthetically

SERVICE CAN BE COMBINED WITH

It can be combined with any service that requires to process large amounts of data (with Spark) or perform machine learning tasks in python (with Jupyter). Examples could be databases, visualization tools, data ingestion tools, etc.

SERVICE OFFERED BY

THE DATA CYCLE HUB

MEMBER
ITI Data Cycle Hub

TYPE
DIH

COUNTRY
Spain

ASK FOR MORE INFO ABOUT THIS SERVICE

The Offer: Datasets in EUHubs4Data

Showing 1 to 20 of 228 entries

FILTERS

- DOMAIN
- COVERAGE
- FORMAT
- PERSONAL DATA PROTECTION
- PUBLIC AVAILABLE URL
- SINGLE DATASET OR REPOSITORY
- LANGUAGES
- MEMBERS

KEYWORDS

Type and search

NAME	DOMAIN	LICENSE TYPE	COVERAGE	DESCRIPTION
1985 AUTO IMPORTS DATABAS	Transport	Creative Commons CCZero (CC0-1.0), Open Data Commo..	United Kingdom	This data set consists of three types of entities: (a) the specif...
A FULL-YEAR DATA REGARDING A SMART BUILDING	Energy	Creative Commons Attribution 4.0 (CC-BY-4.0)	Portugal	The smart building represented by the dataset is divided into fiv...
ACCOMMODATION FACILITIES IN THE VENETO REGION	Tourism	Creative Commons Attribution 4.0 (CC-BY-4.0)	Italy	The content displayed in the table refers to an extraction of the...
ACOUSTIC TEST SET OF THE REVERBERANT MEETING SPEECH IN SLOVAK	Science and Technology	Creative Commons Attribution License 3.0	Slovak Republic	The provided audio samples represent a freely available selected ...
ACTION CITIZEN SCIENCE DATA	Agriculture, Fisheries, Forestry and Food, Environ..	Creative Commons Attribution 4.0 (CC-BY-4.0)	EU-wide, Germany, Italy, Netherlands, Norway, Port..	A collection of datasets created by pollution-focused citizen sci...
AIDA PROJECT DATA	Science and Technology	No license (available on demand)	EU-wide	Heliophysics numerical simulations and high level data
AIR QUALITY DATA STEIERMARK	Environment	Creative Commons Attribution 4.0 (CC-BY-4.0)	Austria	Air quality data Steiermark
AIRLINES DATASET	Transport	Creative Commons CCZero (CC0-1.0), Open Data Commo..	EU-wide	Airlines Dataset Inspired in the regression dataset from Elena Ik...
ALL IRELAND RESEARCH OBSERVATORY (AIRO)	Economy and Finance, Population and Society, Regio..	Private/Proprietary	Ireland	The AIRO Data Store has been created by the team at AIRO and NIRS...
ALTE VIE TRAILS OF THE VENETO DOLOMITES	Education, Culture and Sport, Environment, Tourism	Creative Commons Attribution 4.0 (CC-BY-4.0)	Italy	List of Trails - High Routes of the Veneto Dolomites. The list a...
AMBISONIC ROOM IMPULSE RESPONSES	Sound	Creative Commons Attribution Share-Alike 4.0 (CC-B..	Germany, Spain, United Kingdom	Datasets of Ambisonic room impulse responses from 3 european muse...
ARAGON OPEN DATA	Agriculture, Fisheries, Forestry and Food, Economy..	Creative Commons Attribution 4.0 (CC-BY-4.0)	Spain	Aragón Open Data is the open data portal of the Government of Ara...
ARSO	Environment	Source must be mentioned (ARSO)	Slovenia	Environmental agency of Slovenia, meteorological data
ARTIFICIAL CHARACTERS DATABASE	Science and Technology	Creative Commons CCZero (CC0-1.0), Open Data Commo..	Germany	This database has been artificially generated. It describes the s...

DATASETS

WOOD ANOMALY DETECTION ONE CLASS CLASSIFICATION

LICENCE
Creative Commons Attribution 4.0 (CC-BY-4.0)

DOMAIN
Industry

COVERAGE
EU-wide

FORMATS
PNG

PERSONAL DATA PROTECTION
No personal data

* Please note that the classification is taken from the original source

Wood anomaly detection learning exclusively from non-abnormal wooden textures

Disclaimer: This data is provided by a third party. The DIH identifying this data has no responsibility for its content. Please check the provided link to the data for license terms and potential usage restrictions. In case personal data is included in the dataset, the third party who provides the dataset is the data controller of such personal data. Please note that if you use the datasets for your own purposes, you become an independent data controller and are solely responsible for your compliance with relevant data protection laws relating to the processing and security of personal data, with particular reference, but not limited to, the provisions of the General Data Protection Regulation (GDPR), as applicable to the personal data included in the data.

DATA IDENTIFIED/OFFERED BY

MEMBER
ITI Data Cycle Hub

TYPE
DIH

COUNTRY
Spain

GO TO DATA

ASK FOR MORE INFO ABOUT THE DATASETS/DATA REPOSITORY

Anything wrong? REPORT THIS DATASET

ITI-INSTITUTO TECNOLÓGICO DE INFORMÁTICA AND 2 COLLABORATORS · UPDATED 2 YEARS AGO

20 | New Notebook | Download (740 MB)

Wood Anomaly Detection One Class Classification

Wood anomaly detection learning exclusively from non-abnormal wooden textures

Data Card | Code (0) | Discussion (0)

About Dataset

Context

Surface or Texture-based Anomaly detection is a field within industrial quality inspection that is focused on detecting errors related to the texture of the objects, for instance, stains, color mismatch, cracks, etc. This anomalies may lead to a faulty or even dangerous final product depending on the field, making quality inspection necessary. This is commonly performed employing cameras or other sensors that can detect the patterns of the object's texture. Being able to automatize this process is crucial, as manual inspection has proven to be tedious, slow and inaccurate for many industries.

Moreover, wooden textures present high variability in their structure and color making really difficult to "just learn the correct wooden textures" in order to detect anomalies. However, industry users cannot usually compile all the possible anomalies that might occur during production and their appearances as there are potentially infinite possibilities. Furthermore, manual labeling is a rather subjective and time-consuming task. Taking this into account, one-class classification is the most feasible approach to follow in order to address the problem from a practical industrial perspective.

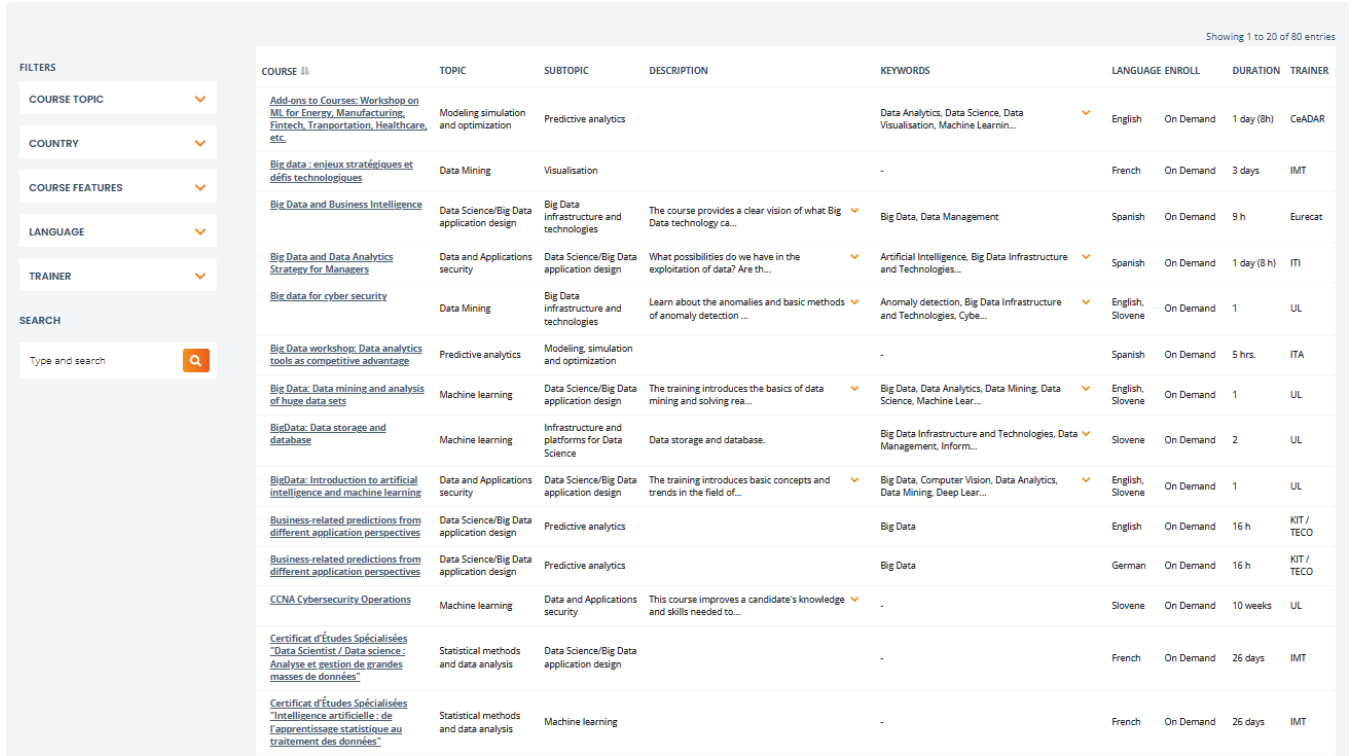
Given this approach, creating a new anomaly detection process only requires to gather a set of objects that will be classified as "correct", and an algorithm will be trained using them. Then, every other object under inspection will be compared to the initial set. Given a significant deviation from the normality, the object can be automatically rejected.

Usability 8.13

License Attribution 4.0 International (CC ...

Expected update frequency Annually

The Offer: Courses in EUHubs4Data



FILTERS

COURSE TOPIC

COUNTRY

COURSE FEATURES

LANGUAGE

TRAINER

SEARCH

Type and search

Showing 1 to 20 of 80 entries

COURSE	TOPIC	SUBTOPIC	DESCRIPTION	KEYWORDS	LANGUAGE	ENROLL	DURATION	TRAINER
Add-ons to Courses: Workshop on ML for Energy, Manufacturing, Fintech, Tranportation, Healthcare, etc.	Modeling simulation and optimization	Predictive analytics		Data Analytics, Data Science, Data Visualisation, Machine Learnin...	English	On Demand	1 day (8h)	CeADAR
Big data : enjeux stratégiques et défis technologiques	Data Mining	Visualisation		-	French	On Demand	3 days	IMT
Big Data and Business Intelligence	Data Science/Big Data application design	Big Data infrastructure and technologies	The course provides a clear vision of what Big Data technology ca...	Big Data, Data Management	Spanish	On Demand	9 h	Eurecat
Big Data and Data Analytics Strategy for Managers	Data and Applications security	Data Science/Big Data application design	What possibilities do we have in the exploitation of data? Are th...	Artificial Intelligence, Big Data Infrastructure and Technologies...	Spanish	On Demand	1 day (8 h)	ITI
Big data for cyber security	Data Mining	Big Data infrastructure and technologies	Learn about the anomalies and basic methods of anomaly detection ...	Anomaly detection, Big Data Infrastructure and Technologies, Cybe...	English, Slovene	On Demand	1	UL
Big Data workshop: Data analytics tools as competitive advantage	Predictive analytics	Modeling, simulation and optimization		-	Spanish	On Demand	5 hrs.	ITA
Big Data: Data mining and analysis of huge data sets	Machine learning	Data Science/Big Data application design	The training introduces the basics of data mining and solving rea...	Big Data, Data Analytics, Data Mining, Data Science, Machine Lear...	English, Slovene	On Demand	1	UL
BigData: Data storage and database	Machine learning	Infrastructure and platforms for Data Science	Data storage and database.	Big Data Infrastructure and Technologies, Data Management, Inform...	Slovene	On Demand	2	UL
BigData: Introduction to artificial intelligence and machine learning	Data and Applications security	Data Science/Big Data application design	The training introduces basic concepts and trends in the field of...	Big Data, Computer Vision, Data Analytics, Data Mining, Deep Lear...	English, Slovene	On Demand	1	UL
Business-related predictions from different application perspectives	Data Science/Big Data application design	Predictive analytics		Big Data	English	On Demand	16 h	KIT / TECO
Business-related predictions from different application perspectives	Data Science/Big Data application design	Predictive analytics		Big Data	German	On Demand	16 h	KIT / TECO
CCNA Cybersecurity Operations	Machine learning	Data and Applications security	This course improves a candidate's knowledge and skills needed to...	-	Slovene	On Demand	10 weeks	UL
Certificat d'Études Spécialisées "Data Scientist / Data science : Analyse et gestion de grandes masses de données"	Statistical methods and data analysis	Data Science/Big Data application design		-	French	On Demand	26 days	IMT
Certificat d'Études Spécialisées "Intelligence artificielle : de l'apprentissage statistique au traitement des données"	Statistical methods and data analysis	Machine learning		-	French	On Demand	26 days	IMT

COURSES

BIG DATA AND DATA ANALYTICS STRATEGY FOR MANAGERS

TOPIC
Data and Applications security

SUBTOPIC
Data Science/Big Data application design

LANGUAGE
Spanish

ENROLL
On Demand

DURATION
1 day (8 h)

TRAINER
ITI

What possibilities do we have in the exploitation of data? Are there currently success stories? What Artificial Intelligence techniques are involved? What is the most appropriate stack of technologies in each case? What options do I have for mass data ingestion and processing? What is the cost of these technologies? What is its degree of maturity? Do I have reliable enough free software options? How should I handle the data optimally? How do I analyze the data to extract knowledge? And finally, what team of professionals do I need and with what training?

This course tries to answer these questions, offering a tour of practical cases that can be applied to the context of business development and a clear vision of the different Big Data technologies that intervene in them. It also focuses on examining what strategy should be implemented in the company to embark on digital transformation and data exploitation. Finally, an approach to the world of data analysis is made, providing the necessary skills to approach a Data Science project in an optimal way, exploring the most efficient and commonly used phases and techniques.
Program:
Unit 1. Introduction to Big Data
Unit 2. Big Data Technologies
Unit 3. Big Data Strategy
Unit 4. Introduction to Data Analytics
Unit 5. Evaluate and manage the Quality of the Data
Unit 6. Machine Learning

Keywords:
Artificial Intelligence, Big Data Infrastructure and Technologies, Data Analytics, Predictive analytics

LEARNING OUTCOMES

Solve the main doubts that management teams face when implementing a Big Data strategy in their company. Thus, it covers from a look at practical cases in the field of data exploitation and analytics, going through a review of the main technologies that today populate the Big Data horizon (for massive ingestion, consultation, analysis, visualization ...), to a roadmap for the adoption of these techniques and technologies in our companies.

COURSE OFFERED BY

THE DATA CYCLE HUB

MEMBER
ITI Data Cycle Hub

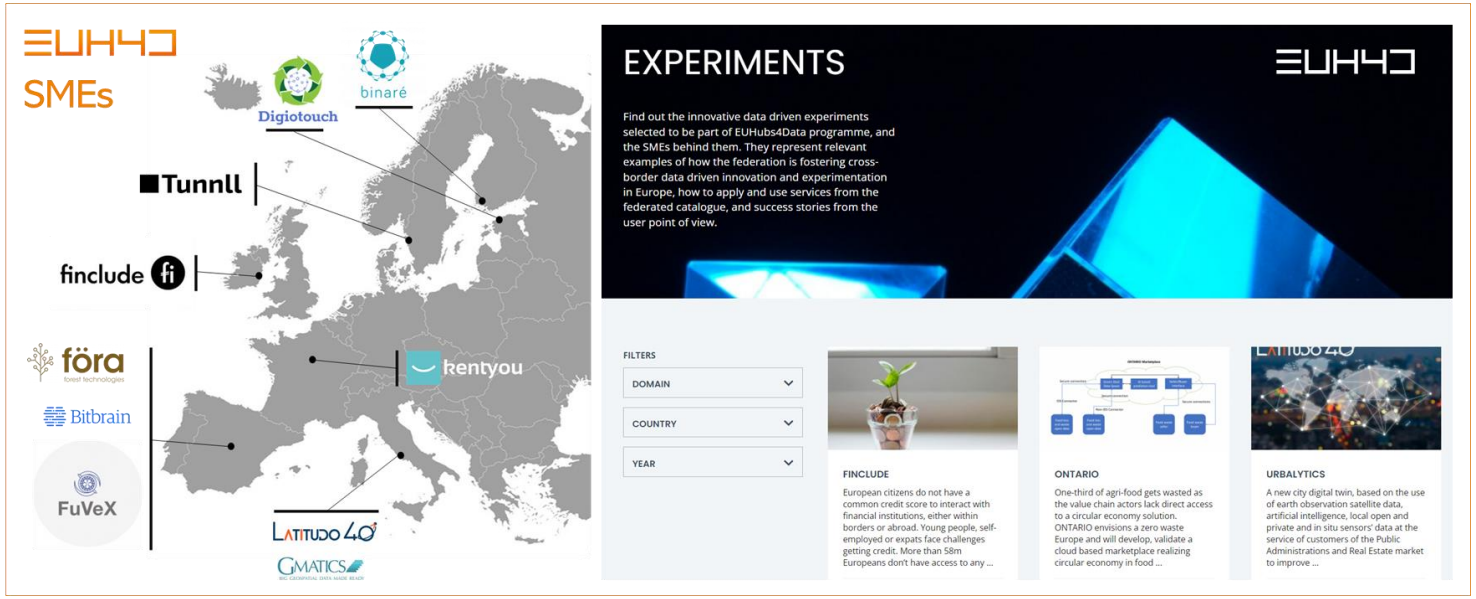
TYPE
DIH

COUNTRY
Spain

ASK FOR MORE INFO ABOUT THIS COURSE

The demand side

- 3 open calls for experiments launched
- 270 applications received
- 42 SMEs selected



- First round of 10 experiments finished (510K Euros successfully delivered)
- Second round of 14 experiments in final review
- Individual exploitation plans
- Third round of 18 experiments started in Feb 2023

Ethics

- Ethics monitoring group created and working
- External Ethics Advisor providing expert advise
- Defined process to assess applications and selected experiments: Ethics toolkit
- Monitoring of experiments
- Assessment of critical ethical issues arising during the project

Visibility

EUH4D
DATAFORUM
Data Spaces Symposium
& Deep-Dive Day
Postillion | The Hague
FEDERATED DATA-DRIVEN EXPERIMENTATION: PRACTICAL LESSONS FROM EUH4D EXPERIENCE
SAVE THE DATE! → 23TH MARCH

DATAWEEK 23
JOIN.LEARN.SHARE.GET VALUE
EVENT:
Data meets Infrastructure at the Edge
→ SAVE THE DATE!
ORGANIZED BY:
BDV BIG DATA VALUE ASSOCIATION
EUH4D
IN COLLABORATION WITH:
RI.SE
13th – 15th June
Luleå, Sweden

EUROPEAN
BIG DATA VALUE FORUM
25-27 OCTOBER 2023 | VALENCIA - SPAIN
IN COLLABORATION WITH
ITI INVESTIGATE TO INNOVATE
INCLIVA | VLC
Biomedical Research Institute
POLITÉCNICA
UNIVERSITAT POLITÈCNICA DE VALÈNCIA
BDV BIG DATA VALUE ASSOCIATION
Data and AI in action: Sustainable impact and future realities

Evolution beyond the project

Vision

The i-Spaces Federation vision is to ***be the reference for experimentation and innovation with industrial, public and personal Data and AI technologies following the European, national and regional values and principles***, combining a **qualified and coordinated** network of Labs/testing facilities and the experience of **responsible practices of data management and AI** ***to support the European companies in their way into the data economy.***

Mission

The Federation's mission is to ***accelerate the evolution and adoption of Data driven innovation and AI Technologies in Europe*** by facilitating **a safe, trustworthy and regulatory compliant environment** for cross-border and cross-sector data-driven experimentation. The Federation links relevant European initiatives on Data and AI in a **single ecosystem** providing a ***sustainable high-quality and global European federated catalogue*** of ***data sources, data-driven services, courses and solutions deployed locally by the i-Spaces***.

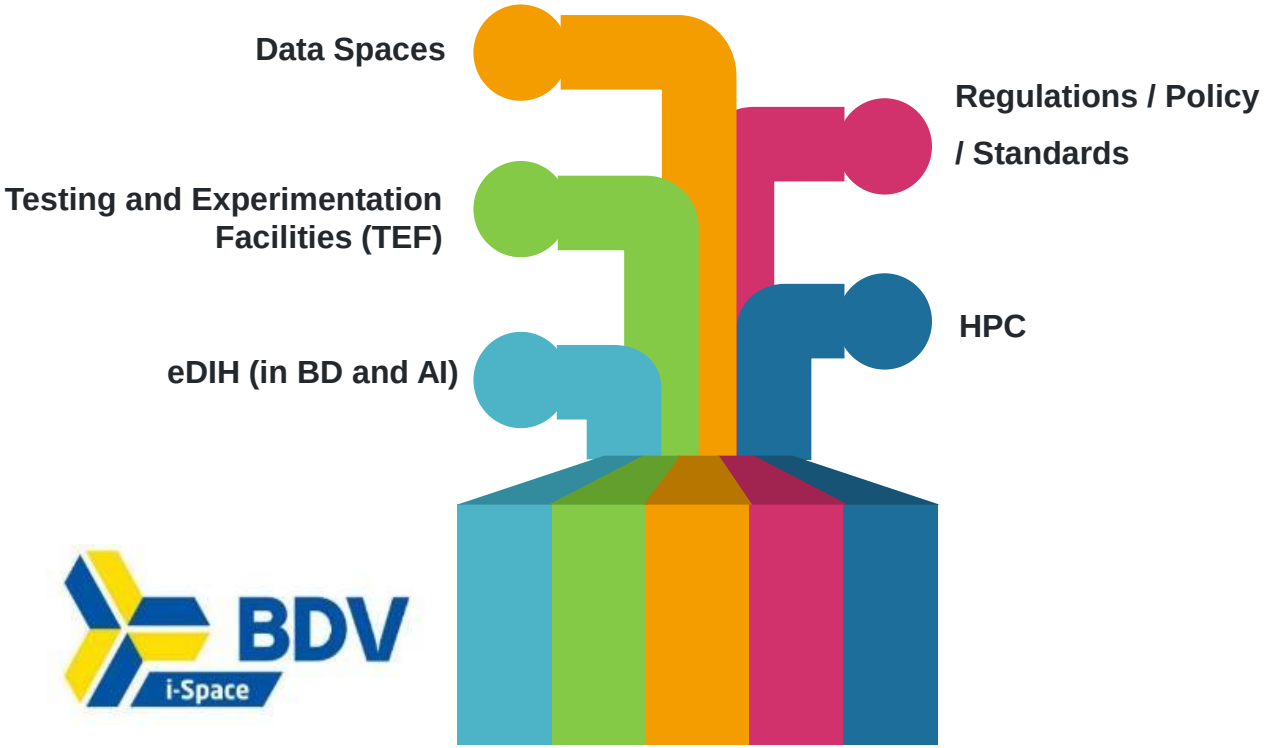
Strategic Objectives (1/2)

- Compose an **integrated and orchestrated network** of i-Spaces working together to a global objective.
- Facilitate the **experimentation and innovation** of the local ecosystem of SMEs and public administration through the local i-Spaces, and **accelerate the adoption of data and AI technologies**, promoting data interoperability, protection and innovation as the core principles of their operation.
- Understand the ***regional and national economy, society and regulations and combine them with the European ones to provide a personalised and quality service to the local ecosystem***.
- Create a **global and qualified catalogue** of data sources, data services and data courses, that brings together and federates the different individual catalogues.

Strategic Objectives (2/2)

- Facilitate the **access of local ecosystems to the main initiatives in Data and AI** in Europe.
- Contribute to the **assessment, definition and validation of current and future European regulations and standards** in data and AI that support the responsible use of data and technology.
- Consolidate the **leadership of Europe in the ethical and human centric use of data and artificial intelligence**, contributing to the international competitiveness.
- Establish **links with the main stakeholders around data spaces**
- Be **sustainable in the future.**

i-Spaces: A key pillar to many European Initiatives/innovative instruments



Why does the i-Spaces have so much potential to generate impact for industries, public services and society?

BDVA i-spaces individually and the i-Space community (and Federation) contribute actively to the following strategies and initiatives:

- European Data Strategy and Data Spaces.
- AI adoption, including AI Testing and Experimentation Facilities (AI TEFs), AI on demand and AI regulatory sandboxes.
- Instruments for the digital transformation, including EDIHs, DTA and network of EDIHs.
- HPC community and EuroHPC JU.

At the same time, i-Spaces and the i-Space community (and/or Federation) contribute to:

- Achieve the objectives of the Digital Decade.
- The adoption of new data and AI regulation and standards.

Value Proposition EUHubs4Data

- *From theory to action*
- *Raise awareness & Prepare the market*
- *Scale up and business arising*

- *Efficiency*
- *Cross sectorial Connection/ Interoperability*
- *Connection to SMEs and reality of data economy*

Accelerate

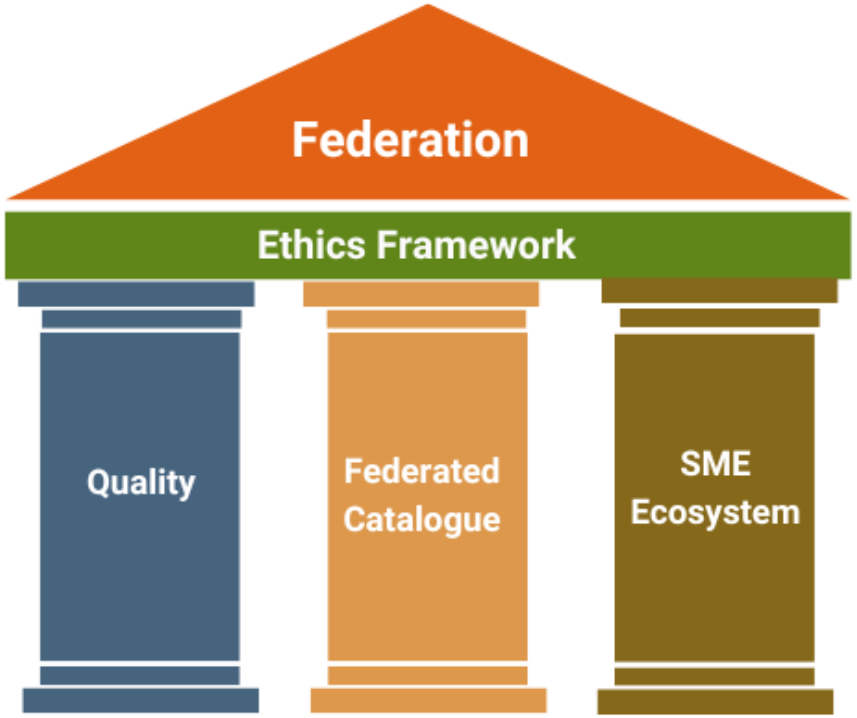
•Technology Lab / point of reference

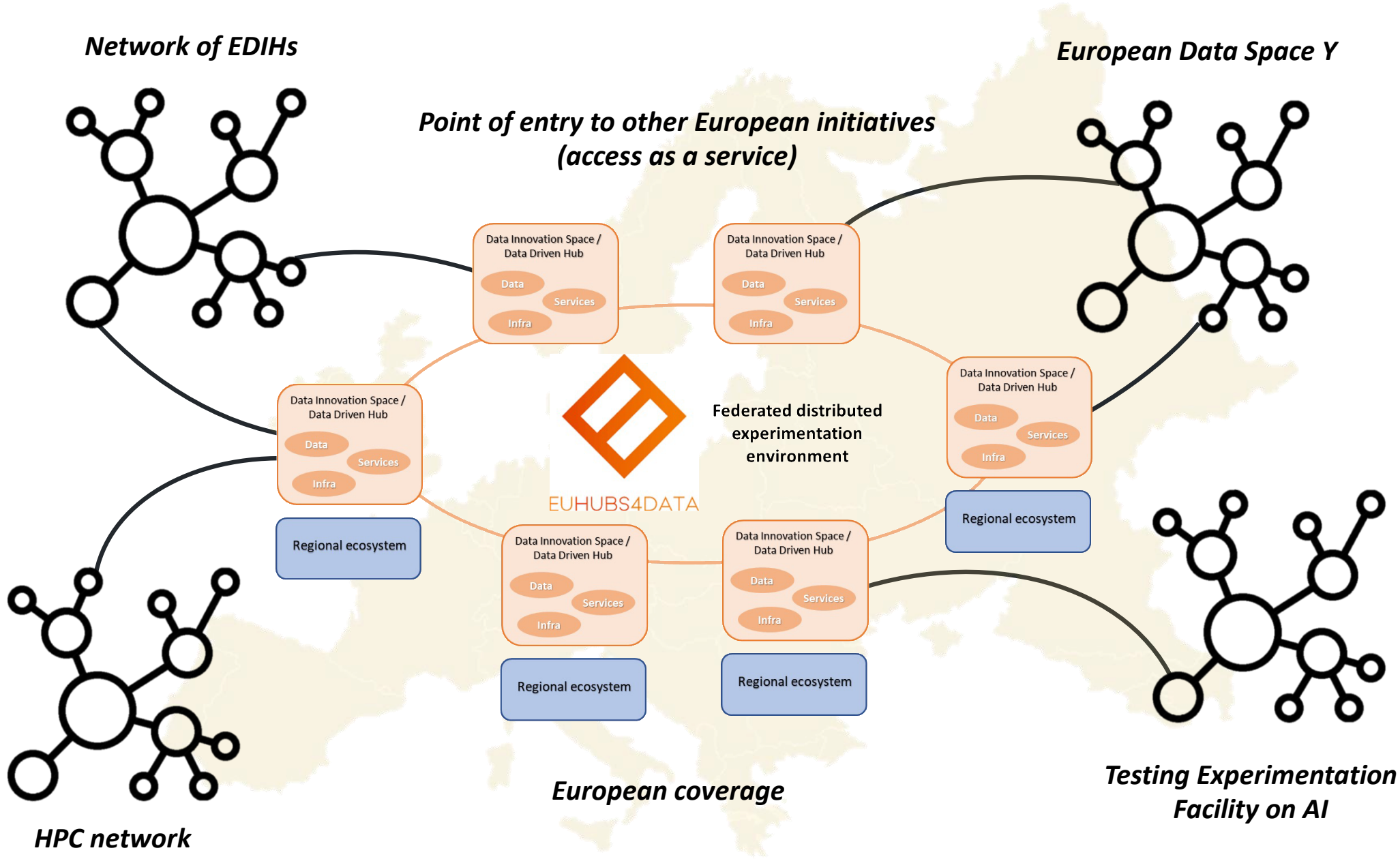
•Sustainability

•Legal & Ethical point of reference

- *Exploit Assets*
- *For other projects*

- *Regulation and standardisation sandboxes*
- *•Ethics and neutrality*




Network of EDIHs
European Data Space Y
Point of entry to other European initiatives
(access as a service)
Data Innovation Space / Data Driven Hub
Data
Services
Infra
Regional ecosystem
EUHUBS4DATA
Federated distributed experimentation environment
HPC network
European coverage
Testing Experimentation Facility on AI

“Alone we can do so little; together we can do so much”

Hellen Keller

DATAWEEK

JOIN.LEARN.SHARE.GET VALUE

Enjoy, learn, share!

EUHUBS4DATA

BDV BIG DATA VALUE ASSOCIATION

European infrastructure and projects:
Panel about the importance to work together in Europe and what is needed on a European level.
How infrastructure and data can secure that we can have a sustainable and prosperous future in Europe?

- Daniel Opalka, EuroHPC Joint Undertaking
- Thor Wikfeldt, EuroCC National Competence Center Sweden (ENCCS), RISE
- Daniel Sáez Domingo, ITI - Instituto Tecnológico de Informática

The user perspective and new users:
Panel on how to reach out to new user how to align the offer to the needs of SME and public authorities.
Share experience and good examples

- Lilit Axner, EuroHPC Joint Undertaking
- Mikael Nyström, Mobilaris
- Sara Json Lindmark, Luleå Science Park
- Apostolos Vasileiadis, EuroCC National Competence Center Sweden (ENCCS)
- Frode Randers, Försäkringskassan

DATAWEEK

JOIN.LEARN.SHARE.GET VALUE

Thank you!

EUHUBS4DATA

BDV BIG DATA VALUE ASSOCIATION

THE DATA IS OUT THERE