



European Health Data Space

Harnessing the power of health data



2 Secondary Use

What are the benefits for whom?

Who will have to make which data available?

How can users apply for access to data?

What are the safeguards?

What infrastructures will provide support?

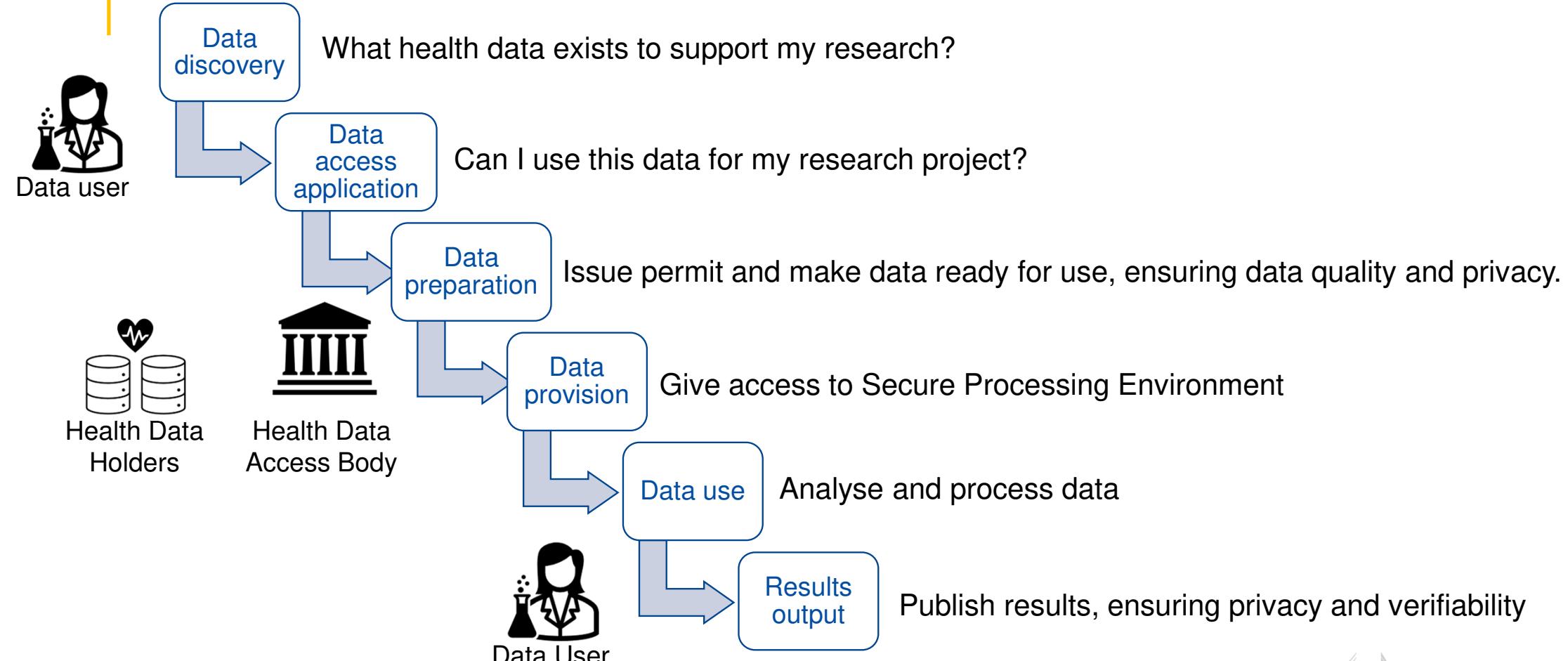
EHDS in a Nutshell – Secondary Use

Secondary use = reuse of data for research and policy making

How?

- Common European rules on who has to make which data available for which purposes and under which conditions
- No need for patient consent to access data, possibility for citizens to opt-out from the reuse of their health data
- Common infrastructure: HealthData@EU
- Data catalogues of available datasets
- Permits for data use, common safeguards

User journey of a researcher



Cross-border secondary use infrastructure

HealthData@EU



Central support services provided by EC



NCP2U - National Contact Point for Secondary Use
EUCP2U - European Contact Point for Secondary Use



Data access services



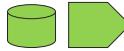
National dataset/metadata catalogue



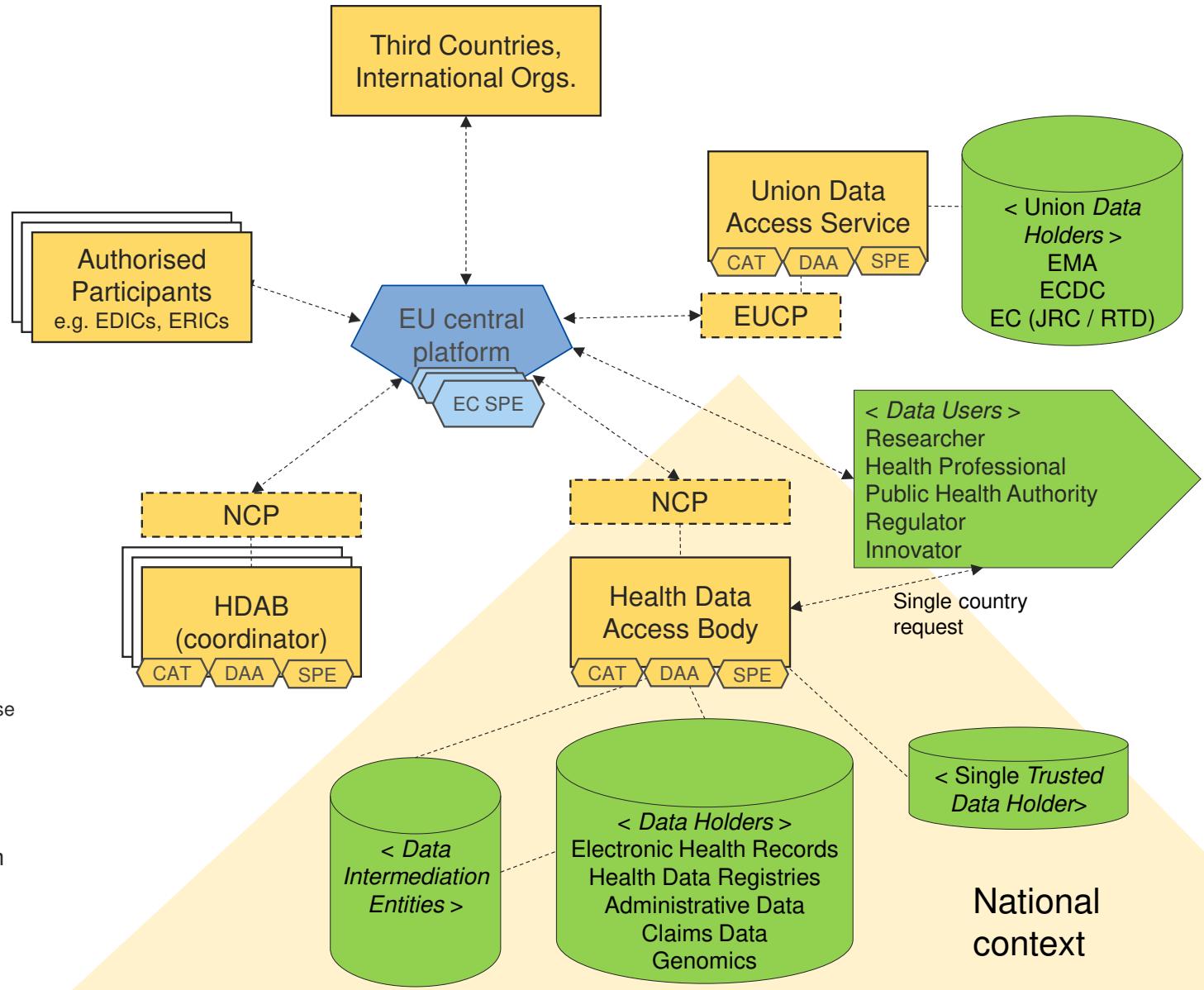
Data Access Application Management system



Secure Processing Environments



Local services provided by/to local partners

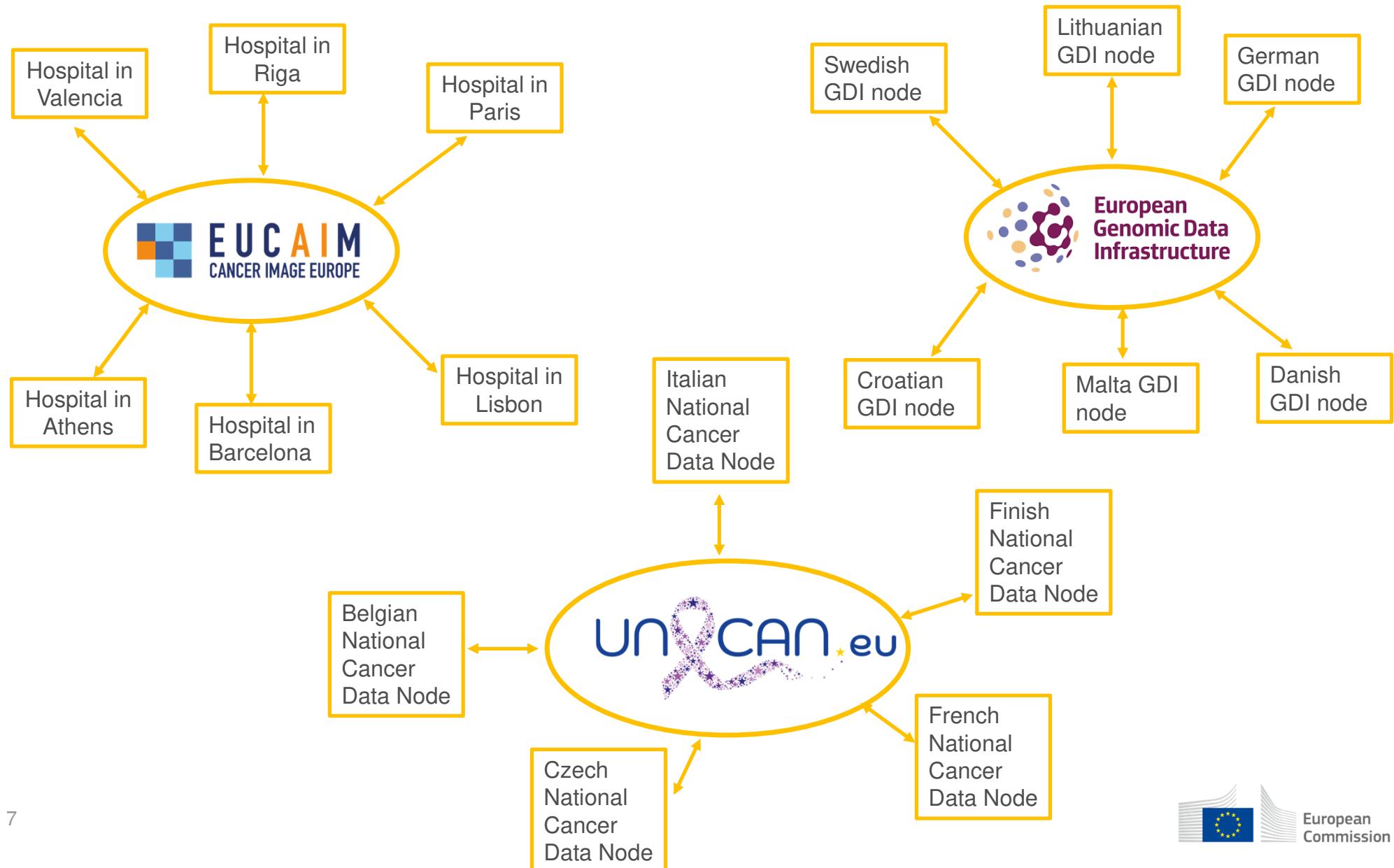


3 Interplay between ERICs/EDICs and the EHDS

How can ERICs and EDICs concretely align with the EHDS?

What could the central platform of these EDICs/ERICs do?

What could the national nodes of these ERICs/EDICs be in the EHDS?



Cross-border secondary use infrastructure

HealthData@EU



Central support services provided by EC

NCP2U - National Contact Point for Secondary Use
 EUCP2U - European Contact Point for Secondary Use

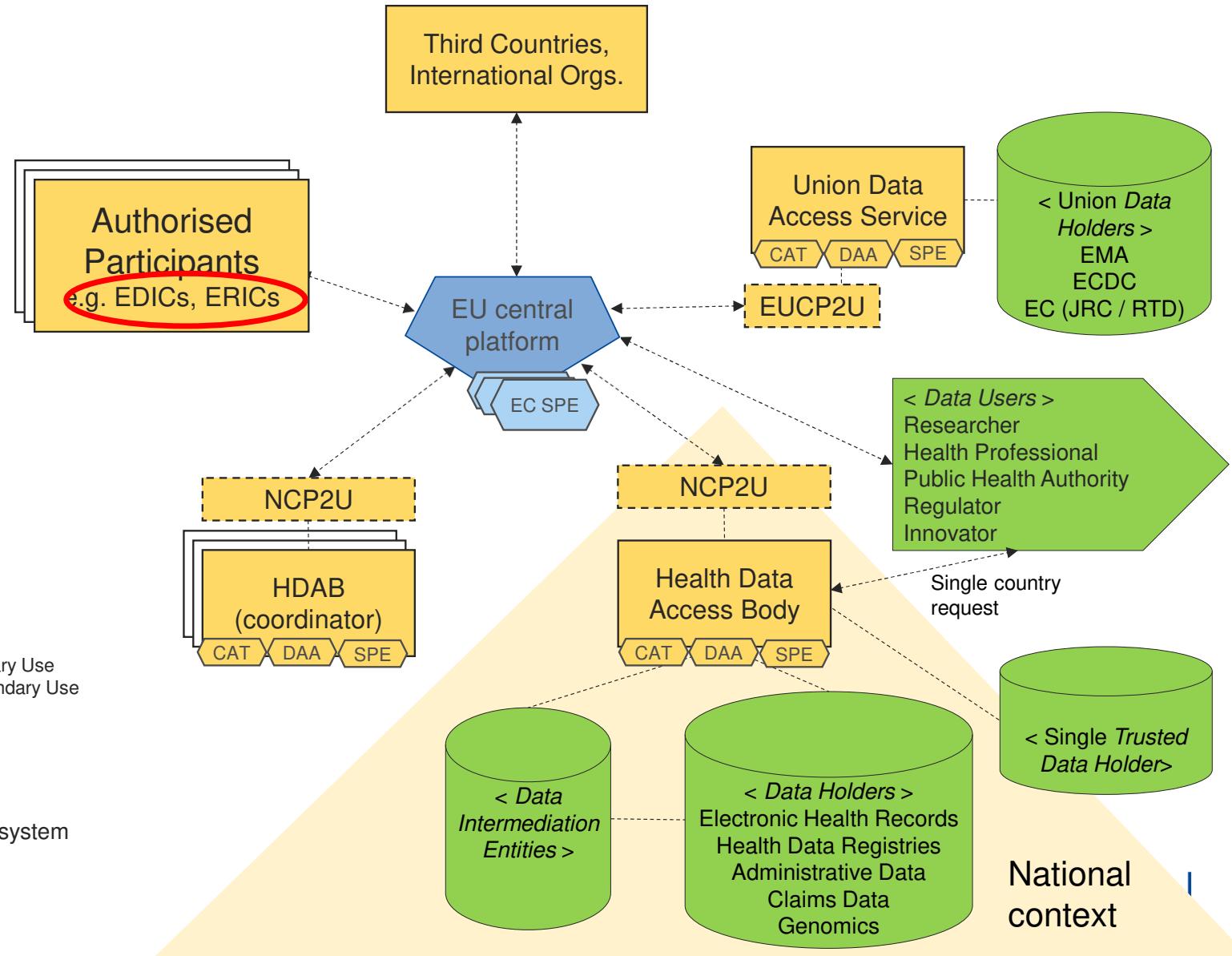
Data access services

CAT National dataset/metadata catalogue

DAA Data Access Application Management system

SPE Secure Processing Environments

Local services provided by/to local partners

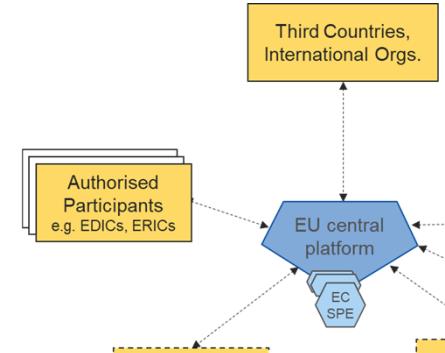


What can the central platforms be?



HealthData@EU bodies

Authorised Participants
Data sharing infrastructures
e.g. EDICs, ERICs



- To be an **Infrastructure established under Union law** (e.g. ERIC, EDIC, ESFRI, EOSC)
- To have the **technical capability to connect to and participate in the HealthData@EU infrastructure**
- Publicly available **metadata catalogue** compliant with the EHDS metadata standard for secondary use
- Be able to **receive and handle data access applications**. They would receive data access applications directly from the EU central platform
- **Decide on data access applications**, authorise and **issue data permits** pursuant to Article 68 to access electronic health data falling within their remit for secondary use and decide on data requests pursuant to Article 69 in accordance with their legal basis and Chapter IV of the EHDS regulation, whenever Authorised Participants are mentioned.

Source of specialised tools, guidelines and SPEs



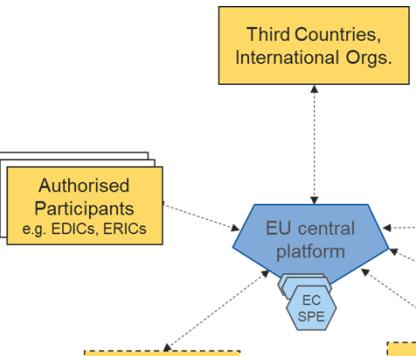
- Hyperontology
- Deidentification guidelines and tools
- AI tools to structure health data
- Tools to curate cancer imaging data
- Specialised SPE



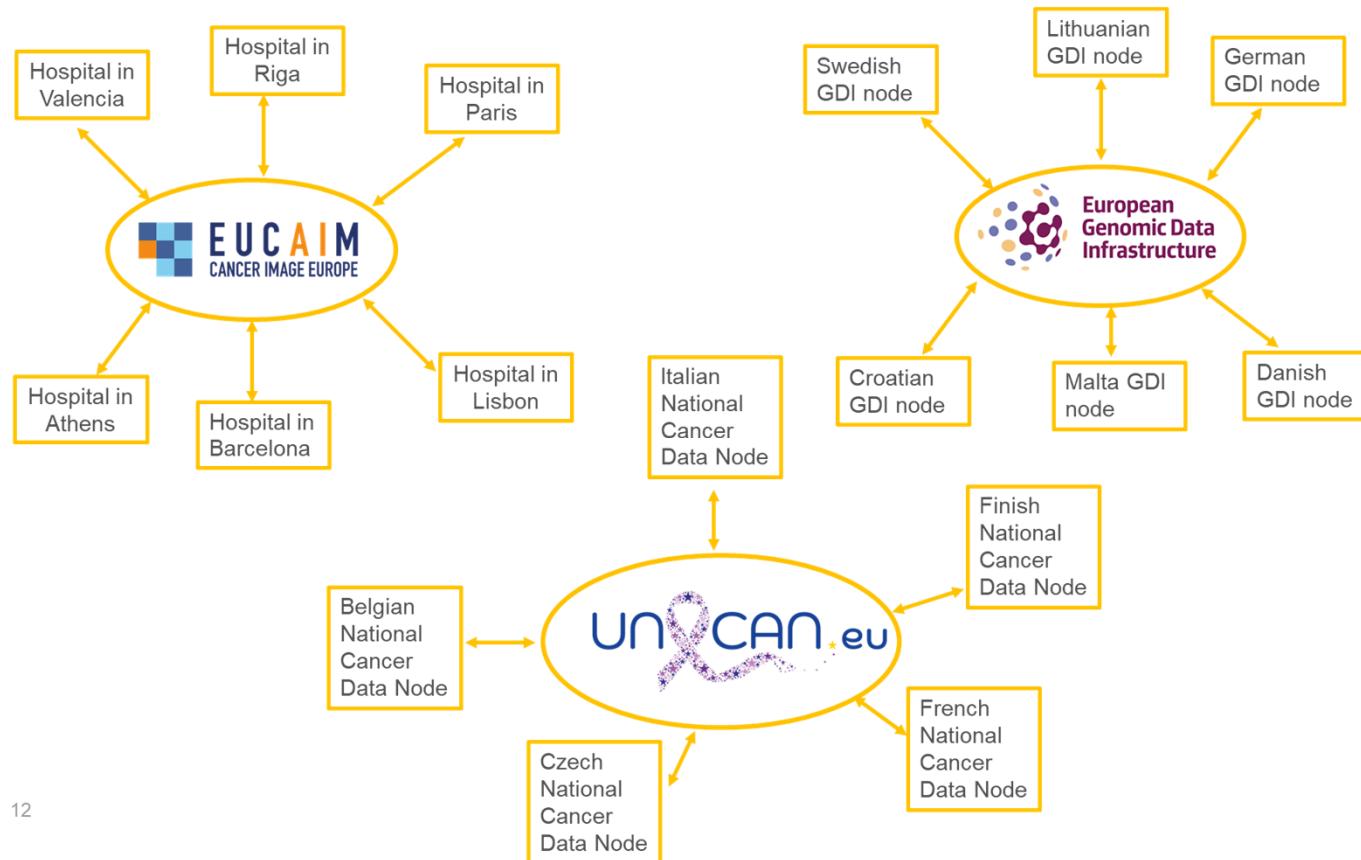
- Data preparation guidelines
- Tools to improve data quality of cancer data at source
- Specialised SPE



- Data preparation guidelines
- BEACON tool to query genomic data
- Guidelines on curation of genomic data at source
- Specialised SPE



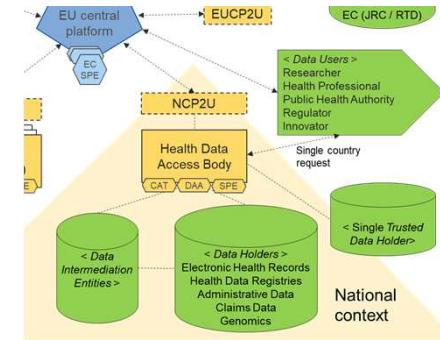
What can the national nodes be?



12

HealthData@EU bodies

Health Data Access Bodies



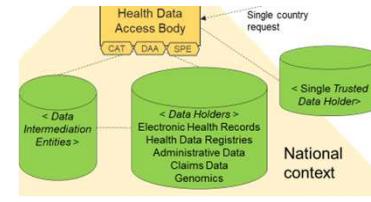
- **Member states designate** one or more HDABs (public sector bodies) to be responsible for making electronic health data available for secondary use.
- **Connected to the HD@EU infrastructure** and the EU central platform
- Hosting and managing the **national metadata catalogue**
- Needs to have human and technical resources to **receive, process and evaluate data access application requests**, and **provision of data permits**.
- Needs to be **in contact with data holders** to receive metadata records from their datasets and ask them to add their datasets in the secure processing environment once a permit was administered.
- Being able **to charge fees** for making electronic health data available for secondary use.
- Provide a possibility for citizens **to opt-out** from the reuse of their data
- Articles 55,57, 58, 59, 63 and 64

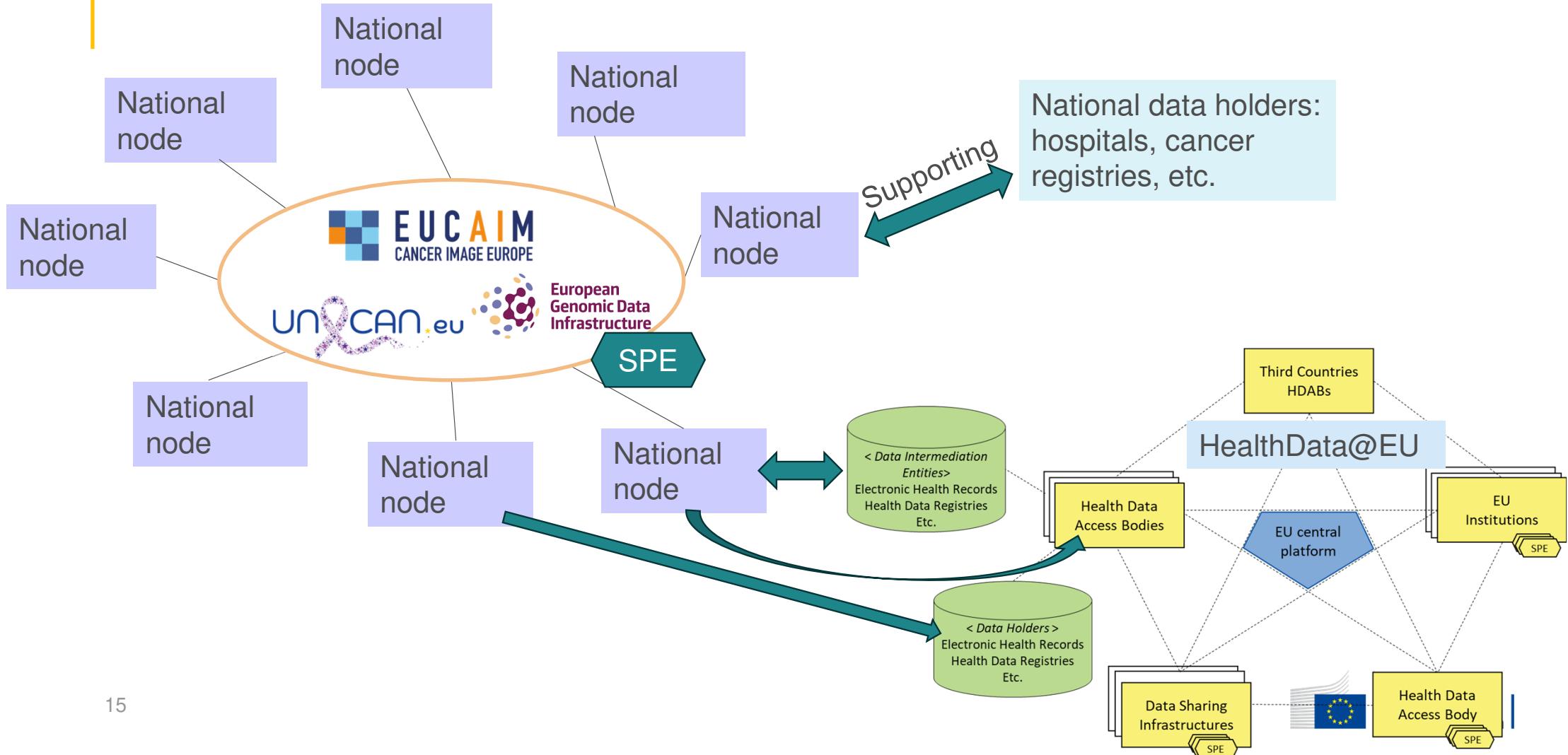
<Health Data Intermediation Entities>

- Be designated by Member States' national law for specific categories of data holders
- Act at national level by providing data from data holders to the relevant HDAB from which the data application form comes
- Do not decide on the permit but would only provide the health data to the HDAB, which would then put these data in a secure processing environment for the data user to access and analyse them.;
- Serve to reduce the administrative burden on holders
- Not connected to HealthData@EU directly, only participates through HDABs
- May not become a trusted data holder

< Single Trusted Data Holder>

- MS may set up procedure for (single) data holder to apply and be designated as a trusted data holder.
- Applications still sent to HDAB, but TDH assesses data access applications that (only) target the data held by it and makes a recommendation to HDAB. Final permit issued by HDAB.
- Must have a secure processing environment at its disposal and provide access to permitted data through it.
- Not connected to HealthData@EU directly, only participates through HDAB.

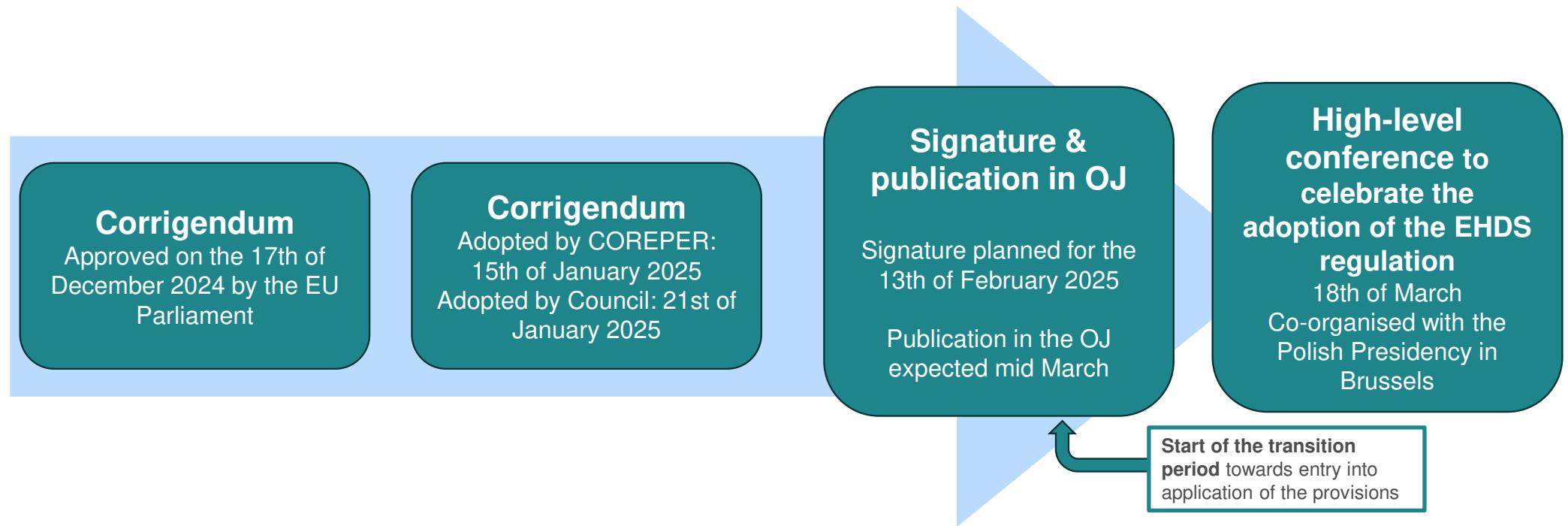




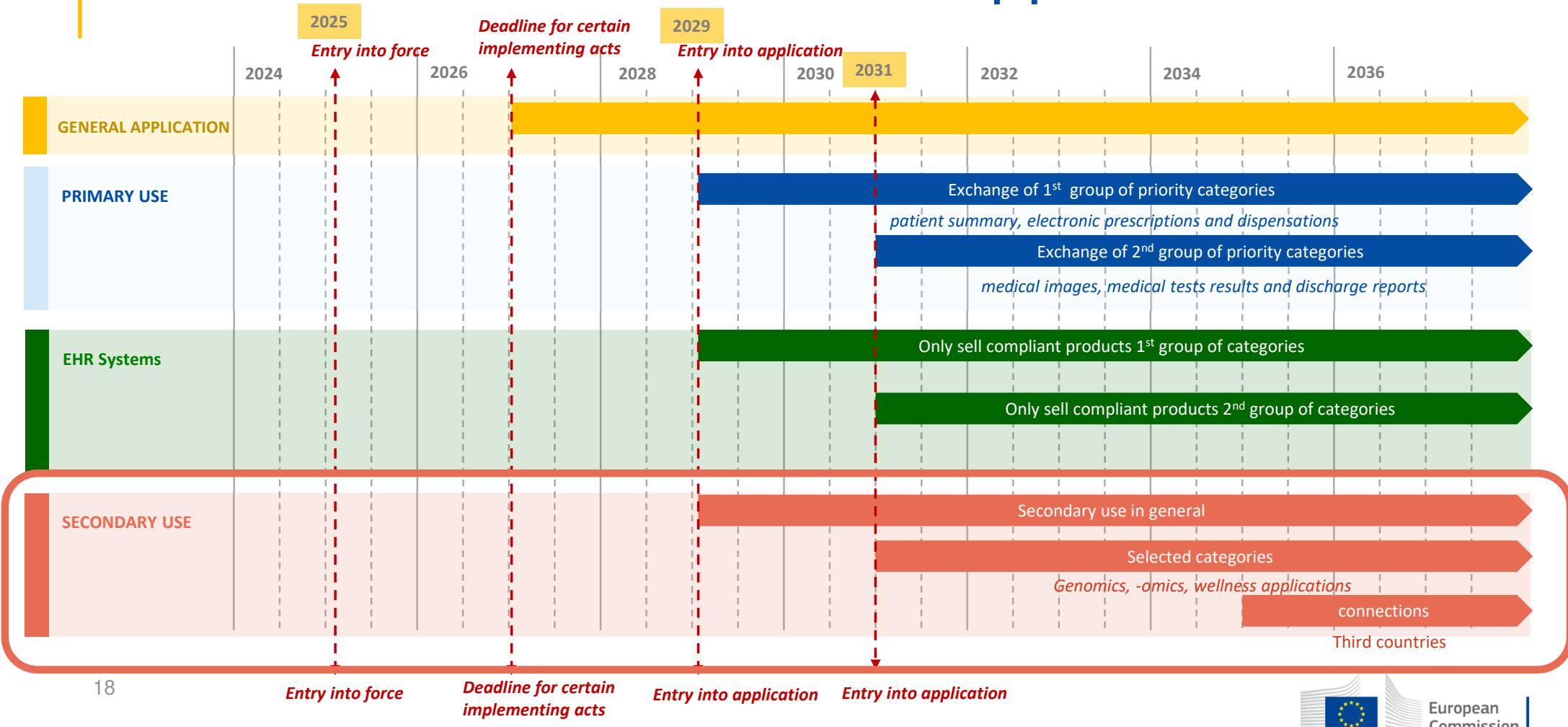
4 Timelines and implementation

By when will all of this happen?
How do we get there?

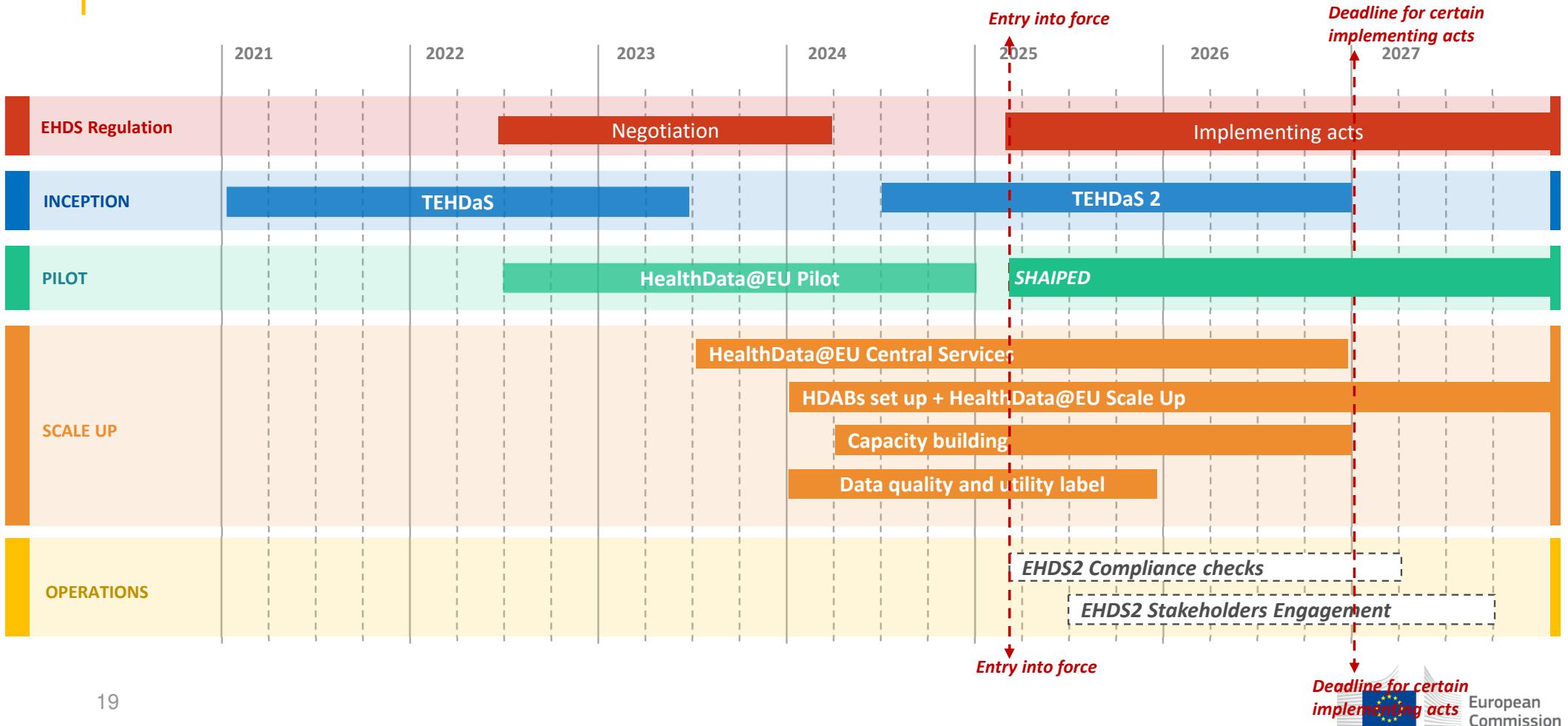
Update on the regulation



EHDS – Overall timeline for application



EHDS2 – Roadmap



Thank you!



HEALTH
DATA HUB



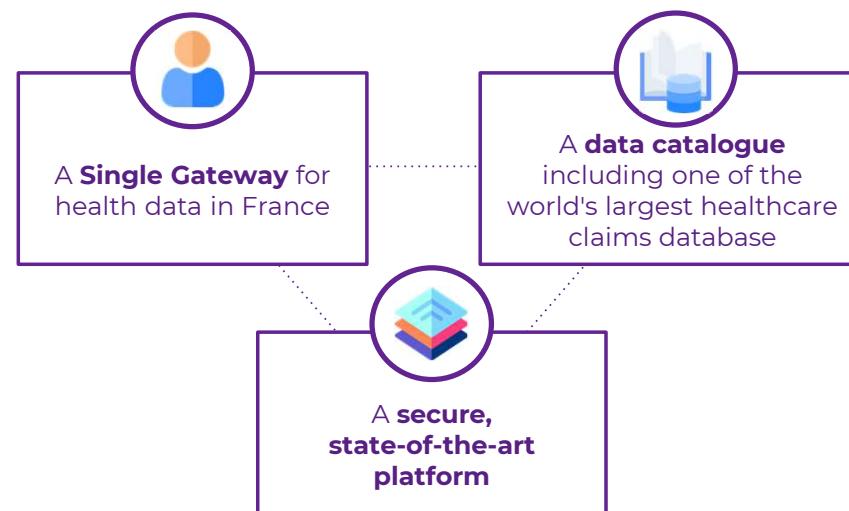
Health Data Hub
[Nom de la
réunion/intervention]

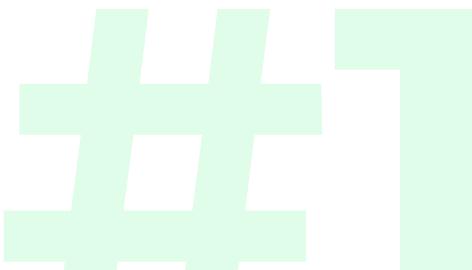
[Lieu (cas échéant)]
[Date]

The Health Data Hub is a French public body, with a role similar to a HDAB

The uses of health data are increasing and it has become essential to ensure access to data sources as quickly as possible.

Created at the end of 2019, the Health Data Hub is a **public body** tasked with **facilitating access to health data for projects in the public interest**, following the granting of open access to the French National Health Data System (Système national des données de santé – SNDS) in 2016.





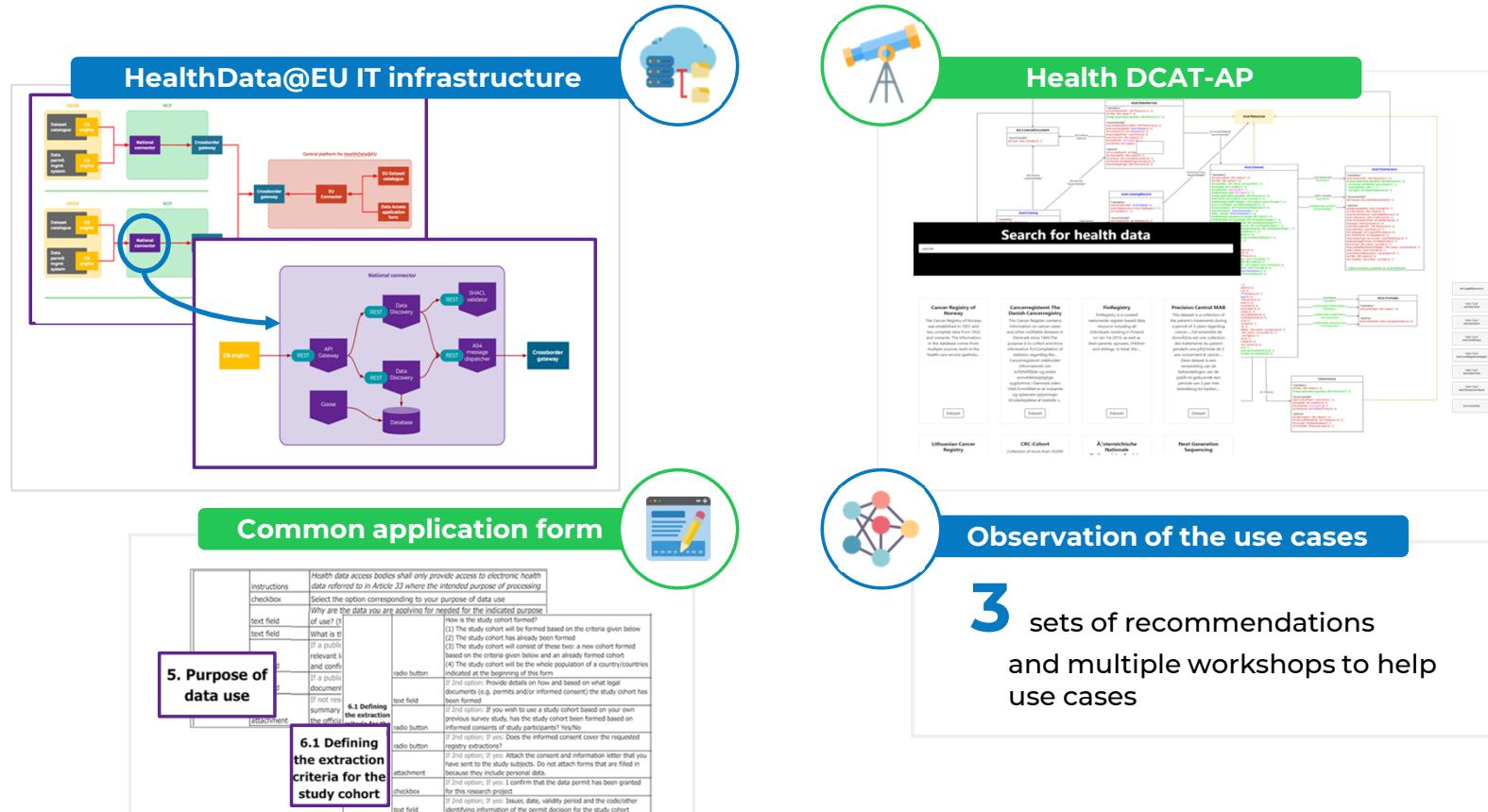
the shaiped project



Intitulé - Date



HD@EU Project Legacy - Key achievements



Multiple related initiatives preparing EHDS regulation implementation (TEHDAS2, QUANTUM, Direct grants, etc.). SHAIPED will continue developing solutions and testing!

SHAIPED's vision is to develop the use of AI in Health through HDABs

3 key objectives:

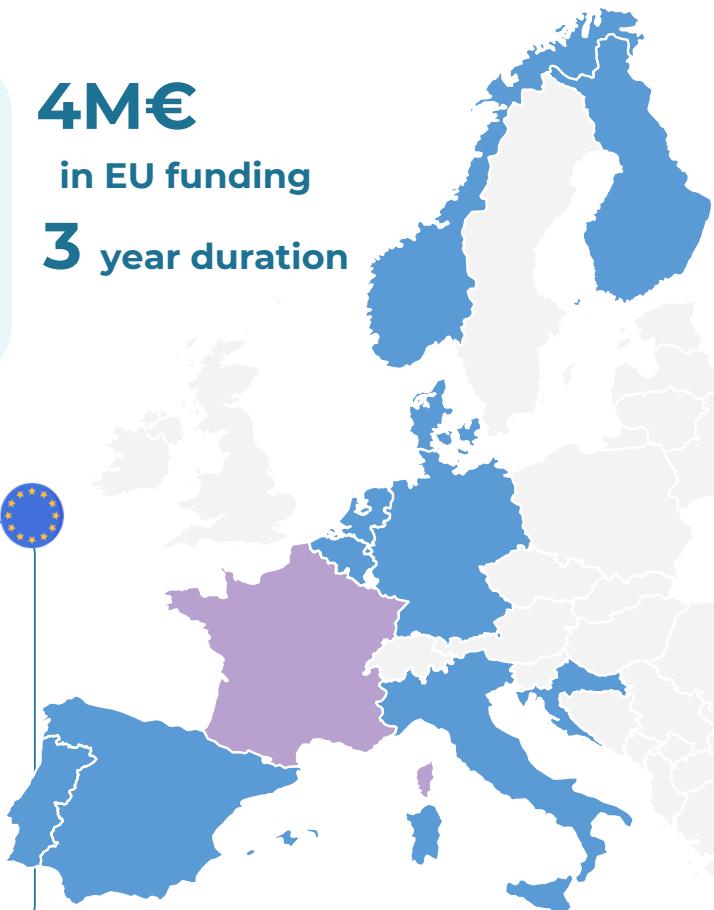
1. Reinforce the capacity of HDABs to provide services in support of AI
2. Foster synergies with the EHDS and the AI Act
3. Conduct concret Use cases

30
Partners
throughout
Europe

11
Member
States
involved



4M€
in EU funding
3 year duration



Intitulé - Date



Objective: Cover the critical steps to integrate and promote AI in the field of Health data



The project's aim is to **promote AI in Healthcare**, to develop the HDABs services needed **for the development, test and deployment of medical devices with AI**.

WP3

Identification of the services, tools and means to reduce the time-to-market of medical devices with AI

- Understanding of AI medical devices life-cycle and description of the HDABs services and tools that can support their development (data access, synthetic data, etc.)
- Analysis of AI Act, EHDS Regulation and of other European legislation related to our topic (GDPR, HTA Regulation, MDR)

WP4

Development of a methodological framework to enhance the deployment perspectives for AI medical devices

- Benchmark of certification process UK/US/EU
- Framework for the use of Real World Data for surveillance and clinical follow-up
- Evaluation : framework to use Real World Data for reimbursement and selection of a technical solutions



WP3 objectives and deliverables



Identification of pathways for HDABs to support AI development and testing, showcase their potential & map legal backdrop



Task 3.1

Define a data lifecycle framework for AI as a Medical Device in the EHDS

Lead: IACS



Task 3.2

Identifying the pathways for HDABs to bolster this framework

Lead: Sciensano



Task 3.3

Showcasing HDABs potential: Synthetic data

Lead: BfArM - HDL



Task 3.4

Regulatory backdrop of AI development, testing and deployment

Lead: UCSC



WP4 objectives and deliverables



Understand current pathways for AI development, testing and deployment, as well as limitations and bottlenecks impeding their effectiveness



Task 4.1

Market access and certification



Task 4.2

Post-Market surveillance & vigilance



Task 4.3

Reimbursement pathways and guidelines for AI as MD selection

3 Use cases to reinforce European competitiveness for AI in Healthcare

#3

Use cases

to test out HDABs
services for AI in
Health



Chronic Kidney Disease Use case (CKD) : Exploration of the performance of an existing algorithm on data from several countries.

Oncology Use case : Set-up a cohort validation to evaluate the performance of AI tools in the detection of metastasis from lung scans imaging and of nodules from mammograms in the framework of a certification process or of a purchase decision.

Cardiology Use case : Evaluation in real world of an AI-based software as a Medical device to manage heart failure, in the framework of reimbursement request.

The Use cases targets identified issues for the deployment of AI in Healthcare, to **reinforce European competitiveness**.



Use case description

Objective

Create a validation cohort of patients to evaluate the performance of AI tools in detecting cancer from a set of standardized annotated images

Description

This UC is a POC to help researchers, clinicians and companies to test AI algorithms for the detection of cancers on a standardized dataset.

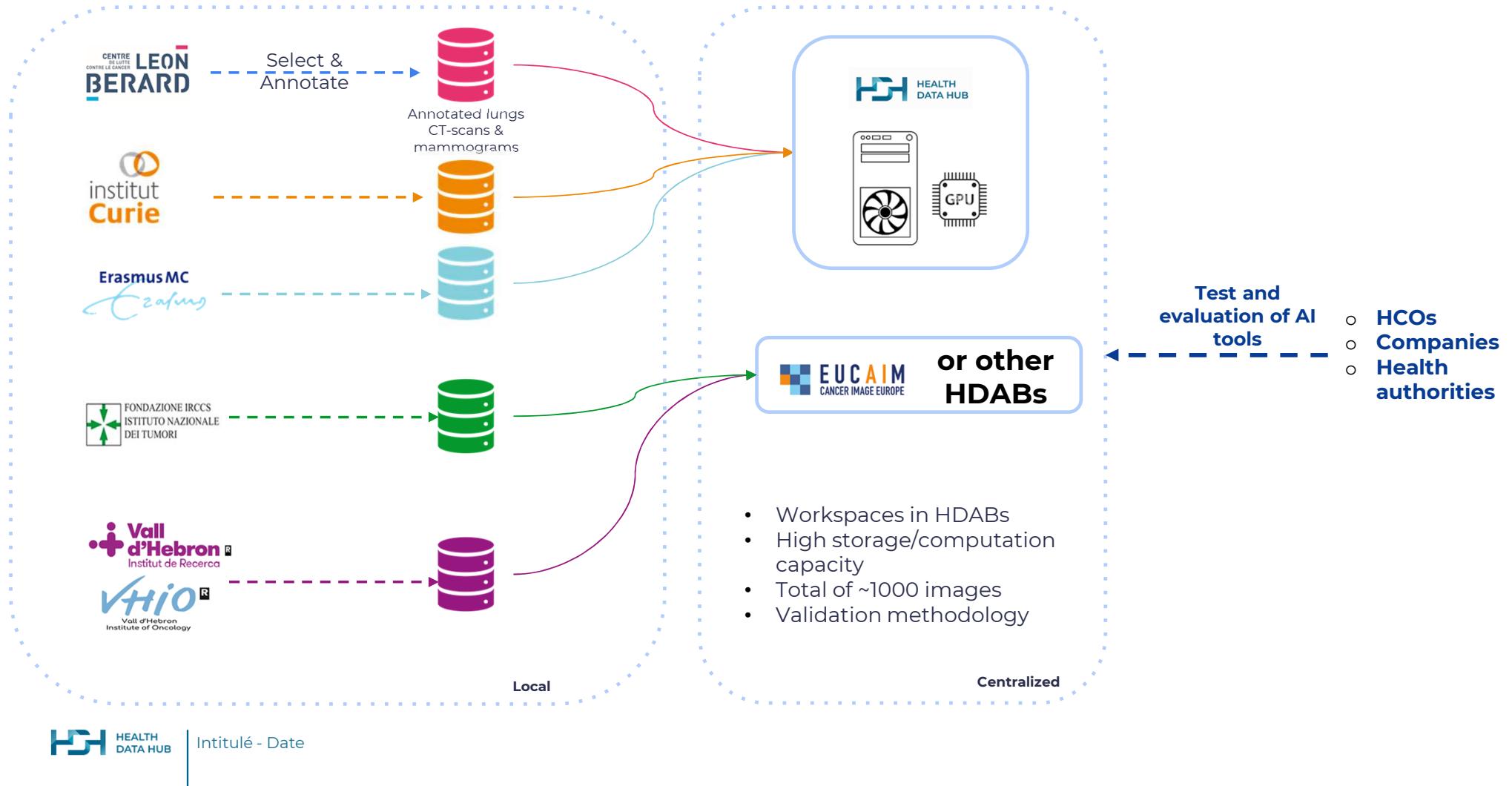
The objective is to build a validation cohort of annotated images from several european cancer centers and to develop a common validation methodology.

The cohort will be made of lung CT-scans and mammograms.

This will help :

- HCOs select AI tools among the different available solutions on the market
- Companies to evaluate the performance of their models before bringing it to the market.
- This could further be used by EMA or reimbursement bodies to evaluate tools as does the FDA

Data flow



Possible avenues for linking SHAIPED & EUCAIM

We are building on existing foundations, and the results of the pilot project provide a solid base upon which we can construct something robust. EUCAIM solutions/tools are also of interest.



SHAIPED focusing on HDAB tools and services, EUCAIM is developing its own set of tools.
EUCAIM tools should be included in the **tool / services mapping**.

EUCAIM legal structure (EDIC?) and role within EHDS (cross-border source of health data) important to understand future functioning with HDABs

Experience and learnings around use cases : Interest to hear about EUCAIM use cases / projects

Oncology use case as a direct avenue for collaboration

What about methodological work (certification, evaluation, post-market surveillance)?



Do you have any questions ?



Suivez-nous sur les réseaux sociaux !



EUCAIM Consortium Meeting

February 5, 2025



- EUCAIM concept
- Sustainable components
- Gaps and complementarities

| | |
|-------------|---|
| 09:00-10:30 | <p>Meeting between EC (DG CNECT and DG SANTE), EUCAIM Project and SHAIPED Project on EHDS-related synergies and collaboration</p> <p>Organised by: EC</p> <p>Participants: DG CNECT, DG SANTE, EUCAIM (UPV, UV, HULAFE), SHAIPED consortium members</p> <p>Format: hybrid</p> <p>Registration for onsite and online participation: https://form.jotform.com/250163120781346</p> <p>Draft agenda:</p> <ul style="list-style-type: none">• Welcome (EC, 5 minutes)• Presentation of the next steps in EHDS Regulation implementation and options for ERIC/EDICs and Research Infrastructures on how to function with the HealthData@EU infrastructure (DG SANTE.C1, 15 minutes)• Presentation of the SHAIPED project to the EUCAIM community (Mario Jendrossek. 15 minutes) |
| | <ul style="list-style-type: none">• Explanation of EUCAIM and potential EUCAIM EDIC to SHAIPED consortium (Ignacio Blanquer. 30 minutes)<ul style="list-style-type: none">- a detailed presentation of the up-to-date EUCAIM service offering in the context of EHDS- opportunities and challenges of future EUCAIM integration into EHDS |
| | <ul style="list-style-type: none">• Discussion on the role EUCAIM can play within the SHAIPED use case on cancer images (25 minutes) |



Co-funded by
the European Union

EUCAIM Project and SHAIPED Project on EHDS-related synergies and collaboration

07/02/2025

EUCAIM service offering

Ignacio Blanquer (UPV)



Co-funded by
the European Union



Data provision models



RWD ENVIRONMENT

(OBSERVATIONAL STUDIES + CANCER SCREENING)

*DATA harvest
model*

*DATA push
model*

ATLAS OF CANCER IMAGES



EUCAIM makes Cancer Image collections newly available to address specific Research and Innovation Questions



Data User defines inclusion and exclusion criteria and data holders generate "on-demand" collections from their available data

Dynamic datasets

EUCAIM integrates and capitalizes existing Cancer Image collections for secondary use



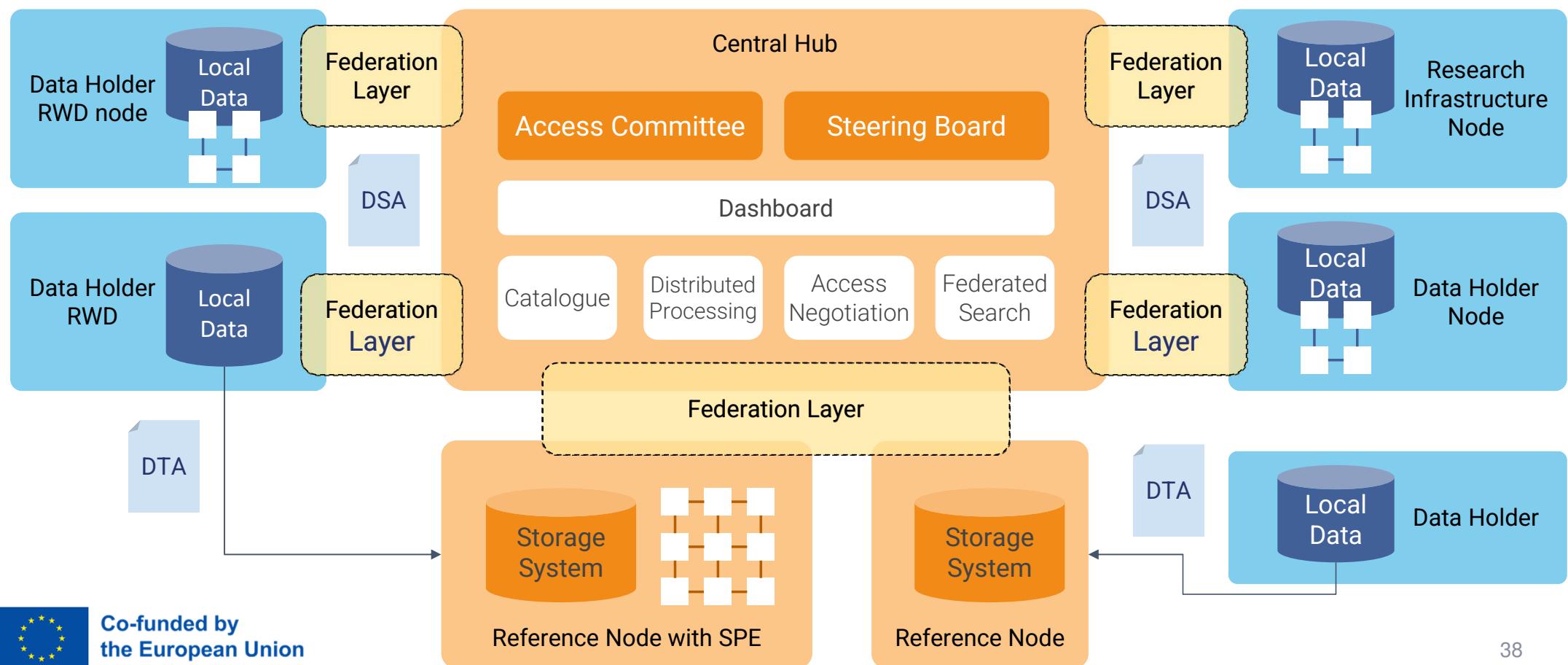
EUCAIM Metadata Catalogue

Data User checks what data is available and selects what better serves them



Co-funded by
the European Union

EUCAIM Federation



Co-funded by
the European Union

Three levels of compliance with the federation



The screenshot shows the EUCAIM Catalogue interface. On the left, there is a sidebar with filters for Providers, Body parts, Imaging modalities, Collection methods, Dataset types, Image access types, Sex, and Vendor. The main area displays a search result for '40 datasets matching the search criteria'. One dataset is highlighted: 'Build an observational study with RWD' (Collection types: Processed datasets, Original datasets, Annotated datasets; Image access: By request; Provider: EUCAIM Observational Studies). Below this, another dataset is shown: 'CHAIMELEON - Lung Cancer Imaging and clinical Data' (Collection types: Annotated datasets, Original datasets; Image access: By request, Restricted access; Provider: Chameleon).

Tier 1: At the Registry level

The screenshot shows the EUCAIM Federated Data Explorer interface. It displays a summary: Provider: 1 / 1, Studies: 12987, Subjects: 10510. Below this, a pie chart shows 'Studies per Collection' with categories: Lung Cancer Training/Validation dataset (October 2024) (blue), Prostate Cancer Training/Validation dataset (October 2024) (orange), Lung Cancer Test dataset (October 2024) (green), Colon Cancer Training/Validation dataset (October 2024) (red), and Lung Cancer Test dataset (October 2024) (purple). A table below lists datasets categorized by collection, provider, studies, and subjects.

| Collections | Provider | Studies | Subjects |
|--|-----------|---------|----------|
| Lung Cancer Training/Validation dataset (October 2024) | CHAMELEON | 1088 | 1088 |
| Lung Cancer Test dataset (October 2024) | CHAMELEON | 273 | 273 |
| Prostate Cancer Training/Validation dataset (October 2024) | CHAMELEON | 1149 | 1149 |
| Prostate Cancer Test dataset (October 2024) | CHAMELEON | 288 | 288 |
| Colon Cancer Training/Validation dataset (October 2024) | CHAMELEON | 904 | 904 |

Tier 2: At the data exploration level

The screenshot shows a Jupyter Notebook titled 'dataset-access-guide'. The code cell contains Python code for plotting images from a dataset. The output cell shows a thumbnail image of a histology slide with the caption '484 items'.

```
import matplotlib.pyplot as plt
import pydicom
# Set the path to the DICOM files
file_path = 'path/to/dicom/files'
# Load the DICOM files into a list
files = pydicom.dcmread(file_path)
# Load the first image
image = files[0].pixel_array
# Plot the image
plt.imshow(image, cmap='gray')
plt.show()
```

Tier 3: At the Data Processing level

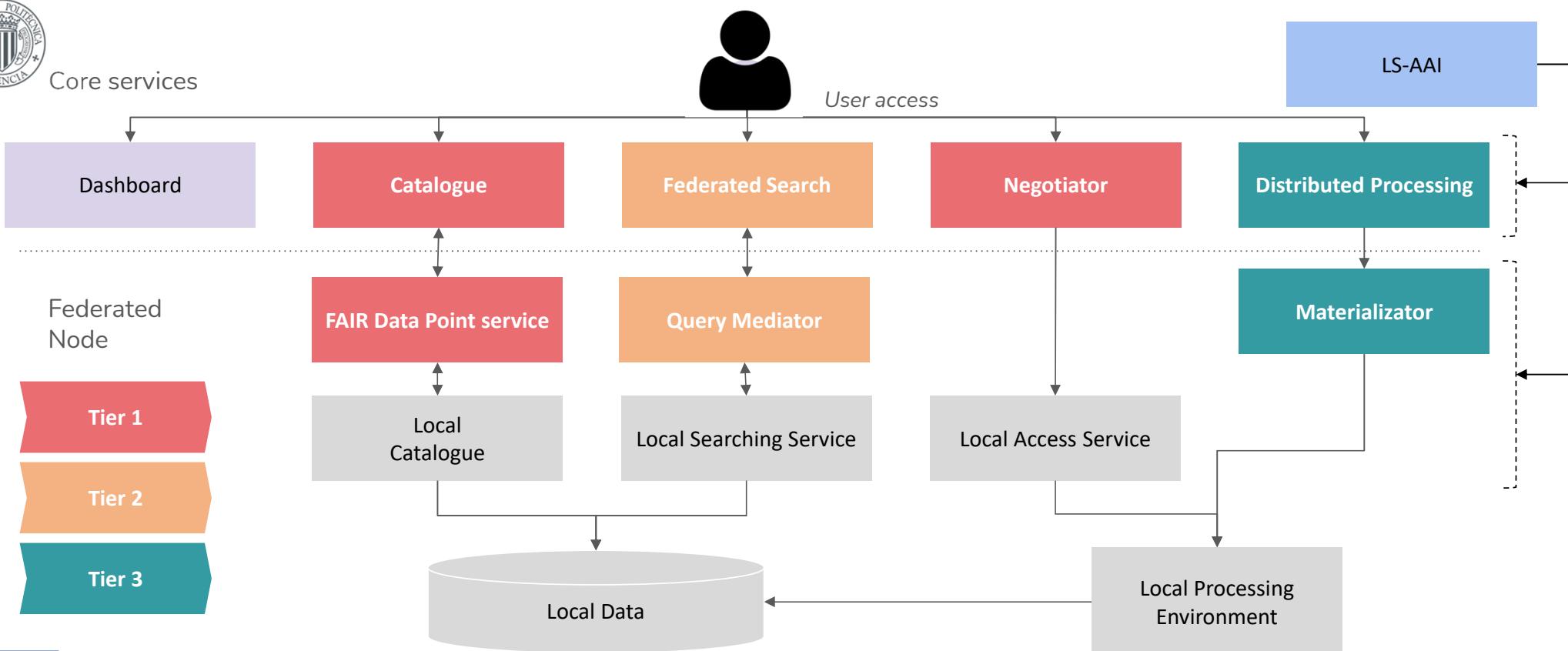


Co-funded by
the European Union

Federation model at technical level



Core services



Co-funded by
the European Union

EUCAIM's Core Services (I)



Catalogue

- A catalogue has 57 datasets registered.
- FAIR Data Point exposing metadata in DCAT.
- Connected to the Access Request Service

The screenshot shows the EUCAIM Catalogue interface. At the top, there are search filters for 'Condition', 'Topography', 'Body part', 'Geography', 'Collection method', 'Order of magnitude', 'Sex', 'Provider', 'Terms of use', and 'Interoperability tier'. Below these, a search bar shows 'Search' and a dropdown menu with 'Dataset series: 40' and 'Dataset(s): 57'. The main area displays search results for 'Build an observational study with RWD' and 'CHAIMELEON - Lung Cancer Imaging and clinical Data'. The 'CHAIMELEON' entry includes a link to 'About this dataset series' and indicates '5 datasets available'. At the bottom, there is a logo for 'Erasmus MC Universitair Medisch Centrum Rotterdam' and 'health RI enabling data driven health & ...'. A small European Union flag is in the bottom left corner.

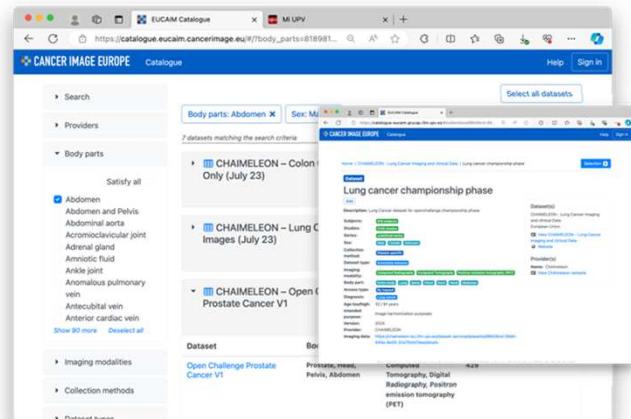
Federated Search

- Provides the list of datasets fulfilling the inclusion criteria and the number of studies.
- 27 searchable items from EUCAIM's CDM.

The screenshot shows the EUCAIM Federated Data Explorer interface. It features a search bar with two queries: 'Modality: Computed Tomography' and 'Modality: Magnetic Resonance Imaging'. Below the search bar, there are sections for 'Patient', 'Clinical Parameters', and 'Image Parameters'. Under 'Image Parameters', there is a 'Modality' section with 'Add all' and a list of modalities: Magnetic Resonance Imaging, Positron Emission Tomography, Single Photon Emission Computed Tomography, and Computed Tomography. To the right, there is a 'Results' summary: 'Provider: 2 / 2', 'Studies: 25490', and 'Subjects: 23235'. Below this, there is a 'Studies per Collection' pie chart and a table of collection details. The table includes columns for 'Collections', 'Provider', 'Studies', and 'Subjects'. The 'Collections' column lists 'ProCancer-I Use Case 1', 'ProCancer-I Use Case 2', 'Lung Cancer Training/Validation dataset (October 2024)', 'Lung Cancer Test dataset (October 2024)', and 'Colon Cancer Training/Validation dataset (October 2024)'. The 'Provider' column shows 'ProCancerI' and 'CHAIMELEON'. The 'Studies' and 'Subjects' columns show various counts for each collection. In the bottom right corner, there is a logo for 'dkfz.'

Registry federation

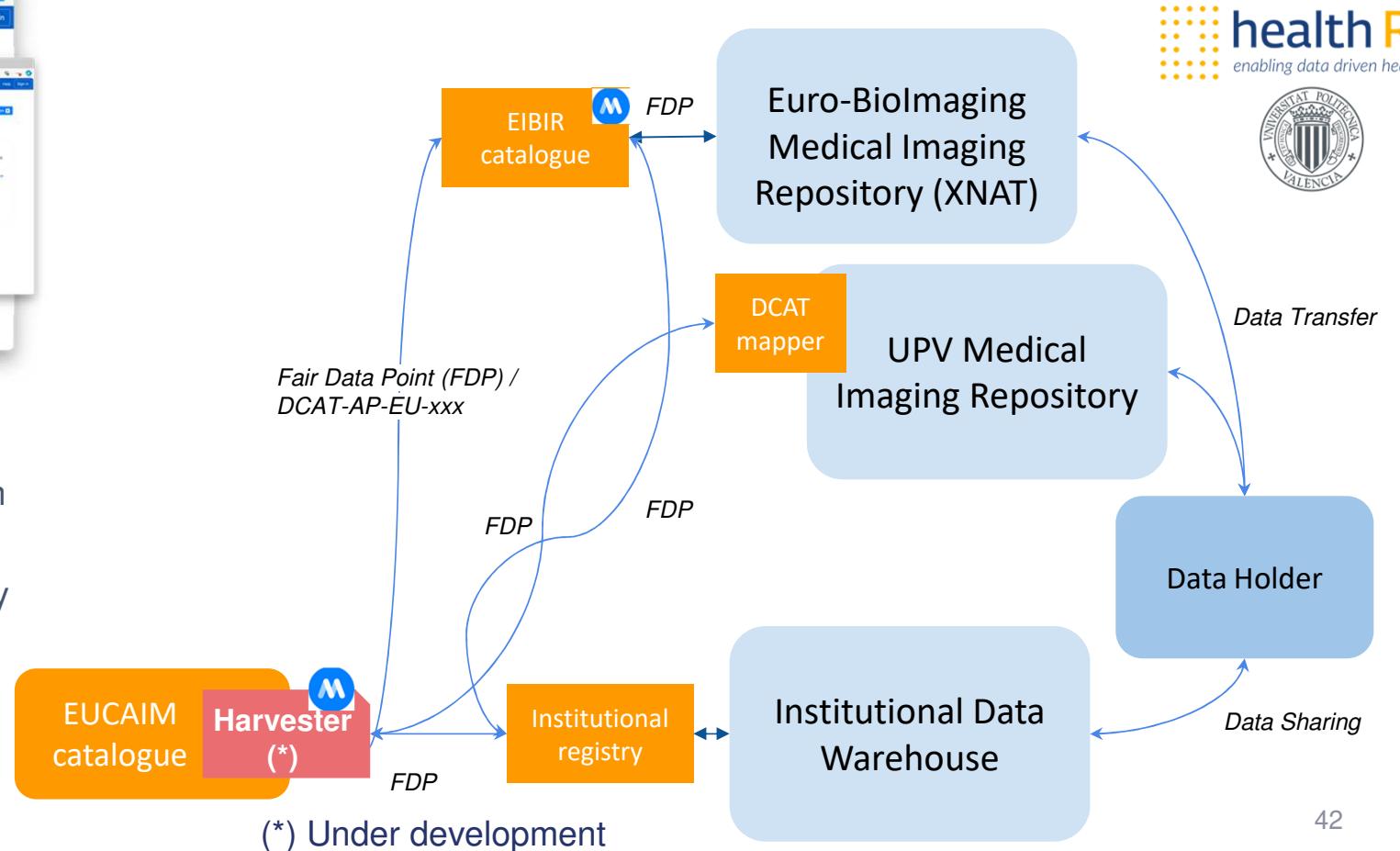
Tier 1



- Register the Datasets metadata in EUCAIM's catalogue
- The Datasets can be organized by centres, data types, etc.
- Following the Health DCAT-AP + EUCAIM extensions.



Co-funded by
the European Union

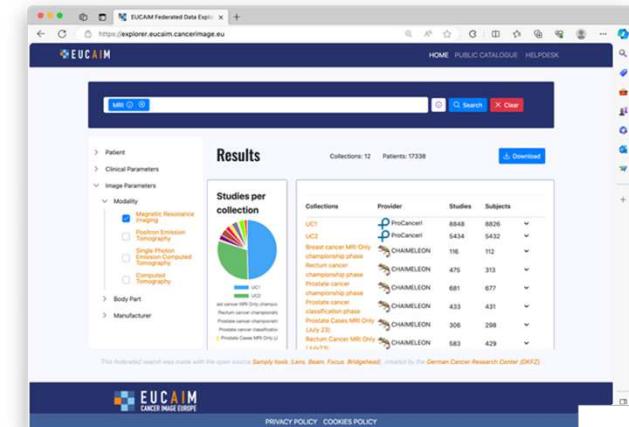
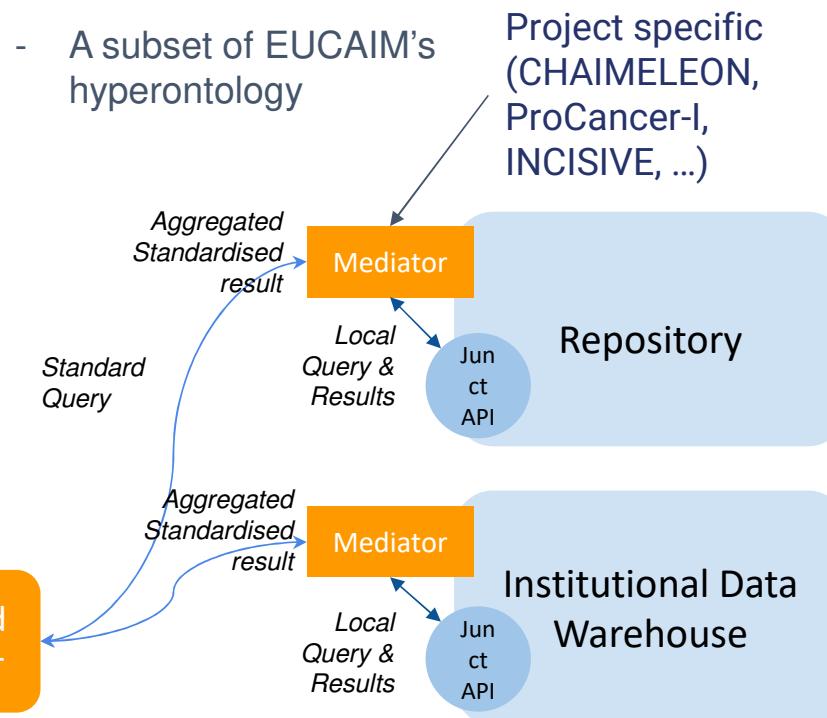


Federated Search

Tier 2



- ✓ Patient
 - > Gender
 - > Age at Diagnosis
- ✓ Clinical Parameters
- ✓ Image Parameters
 - > Modality
 - > Body Part
 - > Manufacturer
- ✓ Tumor Marker Test
 - > Estrogen receptor Ag [Presence] in Breast cancer specimen by Immune stain
 - > Progesterone receptor Ag [Presence] in Breast cancer specimen by Immune stain
 - > Prostate specific Ag [Mass/Volume] in Serum or Plasma
 - > HER2 [Presence] in Breast cancer specimen by Immune stain
 - > Cells.Ki-67 nuclear Ag/100 cells in Breast cancer specimen by Immune stain
- > Histologic Grade
- > Cancer stage
- > Tumor



- Include the datasets in the federated searching
- The searching criterion should be agreed between EUCAIM and the data holders.



Co-funded by
the European Union

Status of Core Services (III)



Access Request Service

- Customised BBMRI-ERIC Negotiator.
- Full life-cycle of dataset access implemented.
- Access forms customised at dataset level.

| Title | Negotiation ID | Created on | Author | Status |
|--|--------------------------------------|------------------|-------------------------|---------------------------|
| My Title | c7de4c05-46b6-4df1-a3dd-d45218163d60 | 2025/01/23 11:23 | Ignacio Blanquer Espert | IN PROGRESS |
| Title | f8a0bb10-29b4-41cb-8a08-06e4d3821853 | 2024/11/27 15:46 | Ignacio Blanquer Espert | UNDER REVIEW |
| CHAIMELEON Lung + breast | 7f40d682-a2dd-4807-94d3-9857df527c5 | 2024/11/21 08:31 | Andrei Stefan Alic | CONCLUDED |
| CHAIMELEON - Rectum Cancer Imaging and clinical Data - Res Available and IN progress | 821bc66e-fb3d-453f-af8d-25627135204 | 2024/11/21 07:41 | Andrei Stefan Alic | CONCLUDED |
| CHAIMELEON Colon DO NOT APPROVE | 29447d59-451c-428b-8517-2756f2fc99d9 | 2024/11/14 12:20 | Andrei Stefan Alic | IN PROGRESS |
| Test CHAIMELEON lung cancer | 996e1756-881f-4d87-a41d-3633eac9de55 | 2024/11/14 12:15 | Andrei Stefan Alic | |
| Titluo 2 datasets | 096b518b-72ae-41c4-aabc-7a4efafa0ccc | 2024/11/14 12:02 | Ignacio Blanquer Es | |

BBMRI-ERIC logo

the European Union logo

Dashboard

- Entry point for the user.
- Data monitoring and User's Library with the datasets with access granted.

Collection Summary: 57 datasets with a total of 42118 samples

Datasets and Subjects

Image Modalities

Explore our PUBLIC CATALOGUE

Become a User

Become a Software Provider

Become a Data Holder

Become a Stakeholder

Cancer Image Europe is pioneering a pan-European federated infrastructure for cancer images, fueling AI innovations

Universitat Politècnica de València logo

Reference nodes



Euro-BioImaging / Health-RI / EMC

Medical imaging repository based on XNAT

- SPE supported by data materializer tool or whitelisting
- Integrated with core services (Catalogue, Negotiator).
- Imaging data: DICOM or NIFTI; Clinical data: JSON, CSV.
- Currently Tier 1 functionality.
- Plans for next period:
 - o Implement FAIR Data Point and upgrade to Tier 3 functionality.



UPV

Storage with SPE from UPV & QP-Insights (Quibim)

- Integrated with core services (Catalogue, Negotiator, SPE, LS-AAI).
- Imaging data: DICOM (or NIFTI); Clinical data: JSON (or CSV).
- Currently Tier 2 functionality.
- Powered with 5 nodes, 480 cores, 3,75TB RAM and 15 GPU accelerators and an additional server with 100 TB.
- Plans for the next period:



oi3M
Instituto de Instrumentación
para Imagen Molecular

Quibim

Roles in the context of EHDS



Co-funded by
the European Union

Cross-border secondary use infrastructure

HealthData@EU



Central support services provided by EC



NCP2U - National Contact Point for Secondary Use
EUCP2U - European Contact Point for Secondary Use



Data access services



National dataset/metadata catalogue



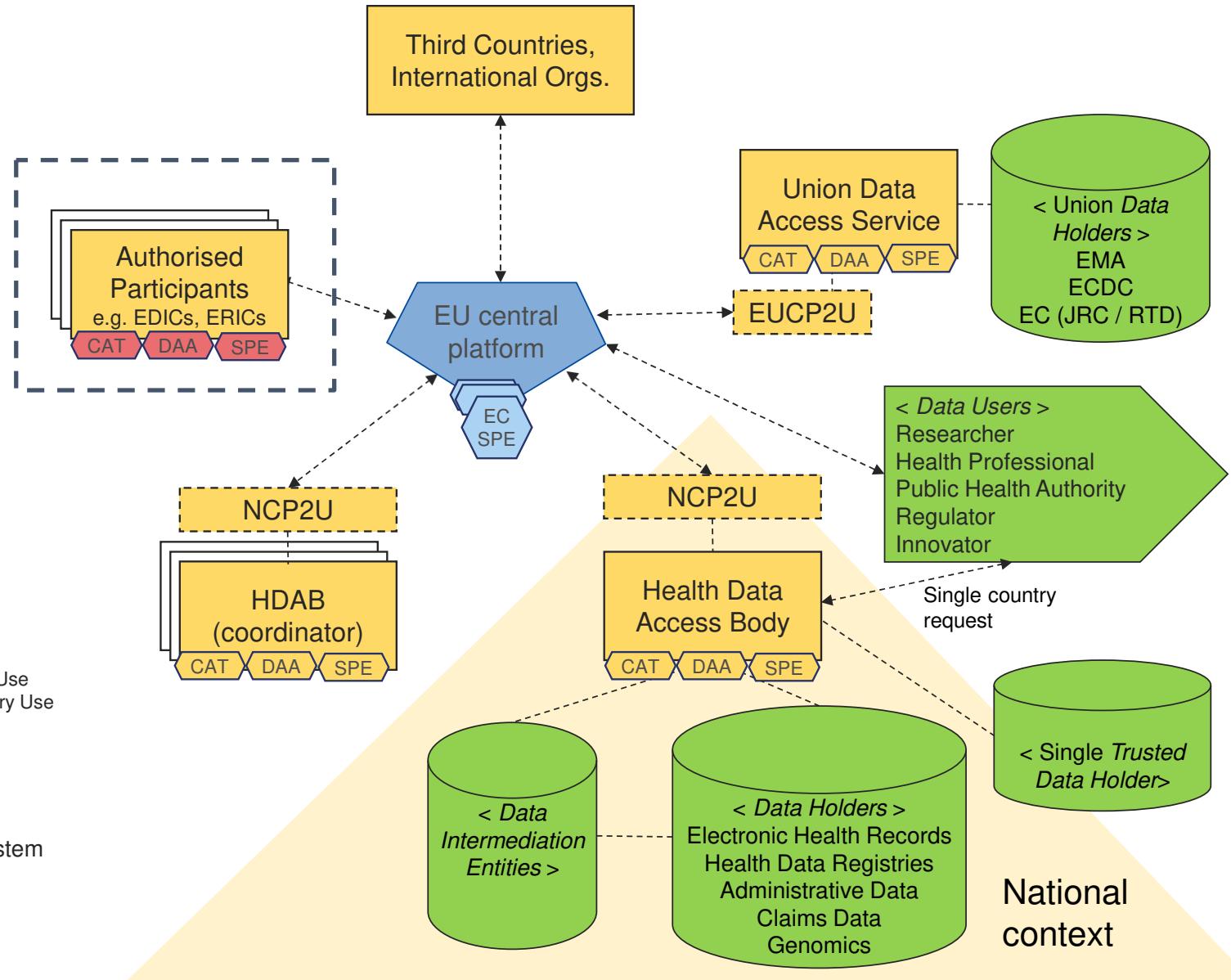
Data Access Application Management system



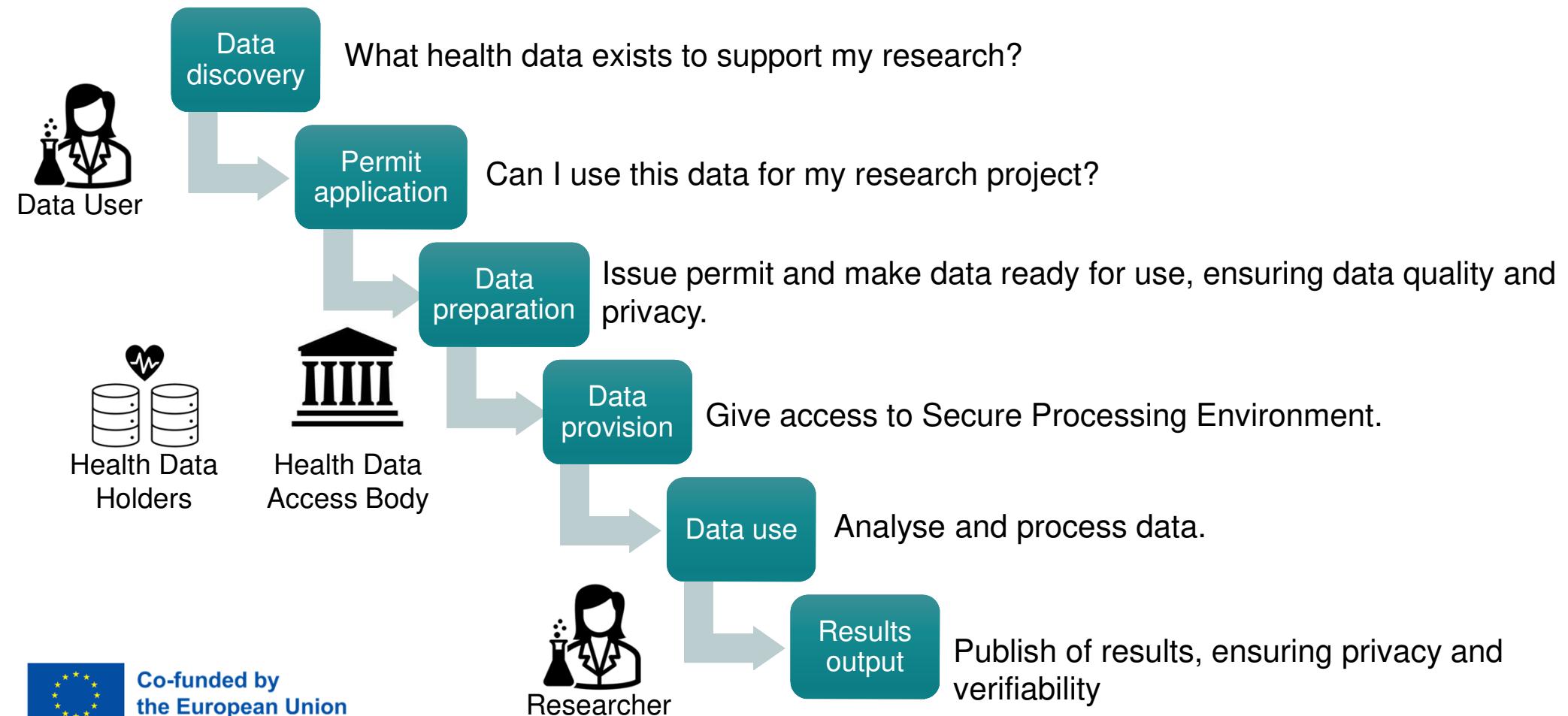
Secure Processing Environments



Local services provided by/to local partners



EHDS Data User Journey



Co-funded by
the European Union

EHDS Data User Journey for the Atlas of Cancer Images



Data User

Data discovery
(catalogue & explorer)

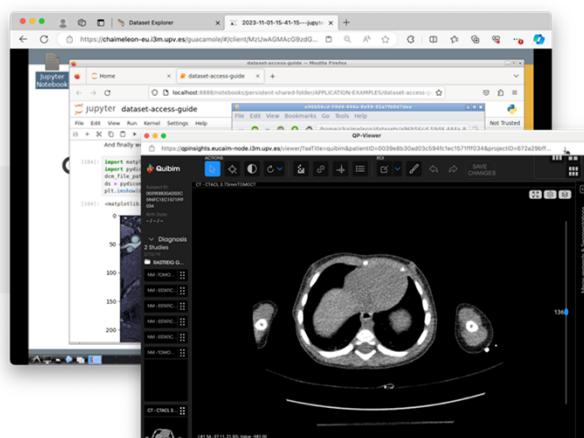
The screenshot shows a web-based interface for data discovery. It features a search bar at the top with filters for 'Modality' (Computed Tomography, Magnetic Resonance Imaging) and 'Type' (Observation, Collection). Below the search bar, there's a summary table with columns for 'Patient', 'Results', 'Provider', 'Studies', and 'Subjects'. A pie chart visualizes the distribution of studies across different providers. On the left, there's a sidebar with navigation links for 'Build an observation', 'Build an observational collection', and 'Image Parameters'.

Data use
(@Reference nodes SPEs or Fed processing)



Researcher

Resource Outputs
(@SPEs or Fed processing)



Data provision
(@Reference nodes)

The screenshot shows the Life Science Identity and Access Management (AAI) interface. It displays a list of users and subgroups under the heading 'EUCAIM #23560'. The interface includes sections for 'Members', 'Subgroups', 'Responsible invitations', and 'Attributes'. A legend on the right side provides definitions for various status icons used throughout the interface.

Permit application
(negotiator)



EUCAIM Access Committee

The screenshot shows the Cancer Image Europe Negotiator interface. It displays a list of negotiations with columns for 'Title', 'Negotiation ID', and 'Created on'. Below this, there's a detailed view of a specific negotiation titled 'CHAMELEON Lung + breast'. An 'Access Form Submission' dialog box is open, showing fields for 'SEARCH PARAMETERS' (e.g., 'Resource ID'), 'RESOURCES (2)' (e.g., 'CHAMELEON - Breast Cancer Imaging and clinical Data'), and 'Access Form' (e.g., 'CHAMELEON - Lung Cancer Imaging and clinical Data').

Data Preparation
(lightweight in the Atlas, negotiator and AAI)



Co-funded by
the European Union

Data Discovery (I)



- EUCAIM's Registry uses the Metadata model defined in EUCAIM
 - Extending the Health DCAT-AP with specific fields for medical imaging and Cancer data.
 - EUCAIM Hyperontology is openly available (<https://zenodo.org/records/14765570>)
 - To be exposed in the registry through FAIR Data Points.
 - To be harvested from the hierarchy of catalogues.

- Willingness to be integrated with national or thematic registries.
 - Beyond EHDS (e.g. EOSC).

The screenshot shows a web browser window for the EUCAIM Catalogue. The URL is https://catalogue-eucaim.grycap.13m.upv.es/Eucaim/eucaim-ui/#/catalogue. The page has a header with the EUCAIM logo and navigation links: Tables, Schema, Up/Download, GraphQL, Settings, Help. A sign-in link 'Hi admin' is also present. The main content area is titled 'Dataset Explorer' with sub-links: Production 3.0.4-beta, Datasets, Models, Documentation, Support. There is a 'Login' button at the bottom right. On the left, there is a sidebar with search filters: Condition (Topography), Body part, Geographical coverage, Collection type, Collection met, Provider, and a search bar. Below this is a section for 'Build an ontology' and 'About this dataset'. It shows '1 dataset available' and a 'Type' dropdown. On the right, there is a table titled 'Dataset Explorer' with columns: Dataset (version) (ID), Flags, Author, Project, Created, Studies, Subjects. The table lists several datasets:

| Dataset (version) (ID) | Flags | Author | Project | Created | Studies | Subjects |
|--|-----------|--------------------|------------|---------------------------|---------|----------|
| Lung Cancer Training/Validation dataset (October 2024) (1.0) | Published | Ana Jimenez-Pastor | CHAIMELEON | 01/11/2024, 21:51:53 CET | 1088 | 1088 |
| (bae8f8fc-b1bc-a46b-9c59-58071de64473) | | | | | | |
| Lung Cancer Test dataset (October 2024) (1.0) | Published | Ana Jimenez-Pastor | CHAIMELEON | 01/11/2024, 21:50:06 CET | 273 | 273 |
| (49f05d44-19c4-4678-acd4-15bbc093d6a00) | | | | | | |
| Prostate Cancer Training/Validation dataset (October 2024) (1.0) | Published | Ana Jimenez-Pastor | CHAIMELEON | 25/10/2024, 13:28:12 CEST | 1149 | 1149 |
| (97c6494a-e42c-47d6-83f7-76cd5865350) | | | | | | |
| Prostate Cancer Test dataset (October 2024) (1.0) | Published | Ana Jimenez-Pastor | CHAIMELEON | 25/10/2024, 13:09:40 CET | 288 | 288 |
| (1.0) | | | | | | |

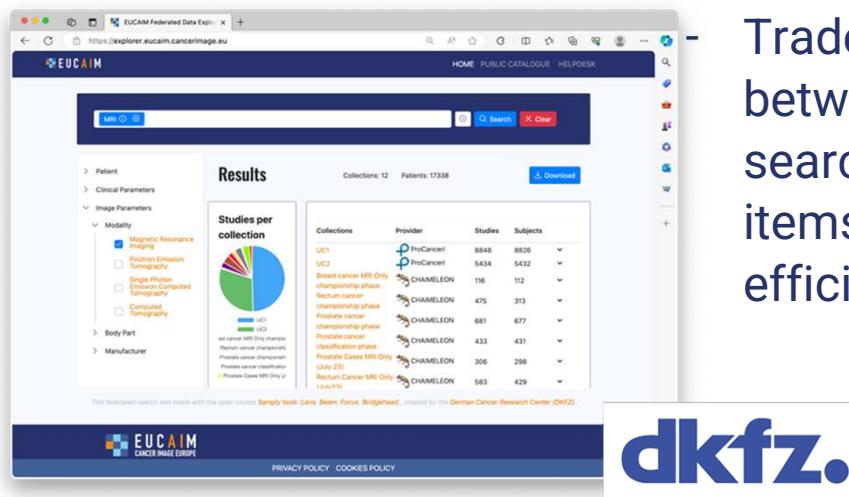


Co-funded by
the European Union

Data Discovery (II)



- The Federated search is a key element on data discovery
 - Data access request and data preparation are time-consuming processes and therefore it is key to understand if a dataset has the appropriate information.
- This can only work on already collected data
 - Valid for intermediate entities and Data holders with Data Warehouses.



- It could be relevant for national registries to have a quicker and more accurate view of the data available
- Trade-off between searchable items and efficiency.

- > Patient
- > Gender
- > Age at Diagnosis
- > Clinical Parameters
- > Image Parameters
 - > Modality
 - > Body Part
 - > Manufacturer
- > Tumor Marker Test
 - > Estrogen receptor Ag [Presence] in Breast cancer specimen by Immune stain
 - > Progesterone receptor Ag [Presence] in Breast cancer specimen by Immune stain
 - > Prostate specific Ag [Mass/Volume] in Serum or Plasma
 - > HER2 [Presence] in Breast cancer specimen by Immune stain
 - > Cells.Ki-67 nuclear Ag/100 cells in Breast cancer specimen by Immune stain
- > Histologic Grade
- > Cancer stage
- > Tumor

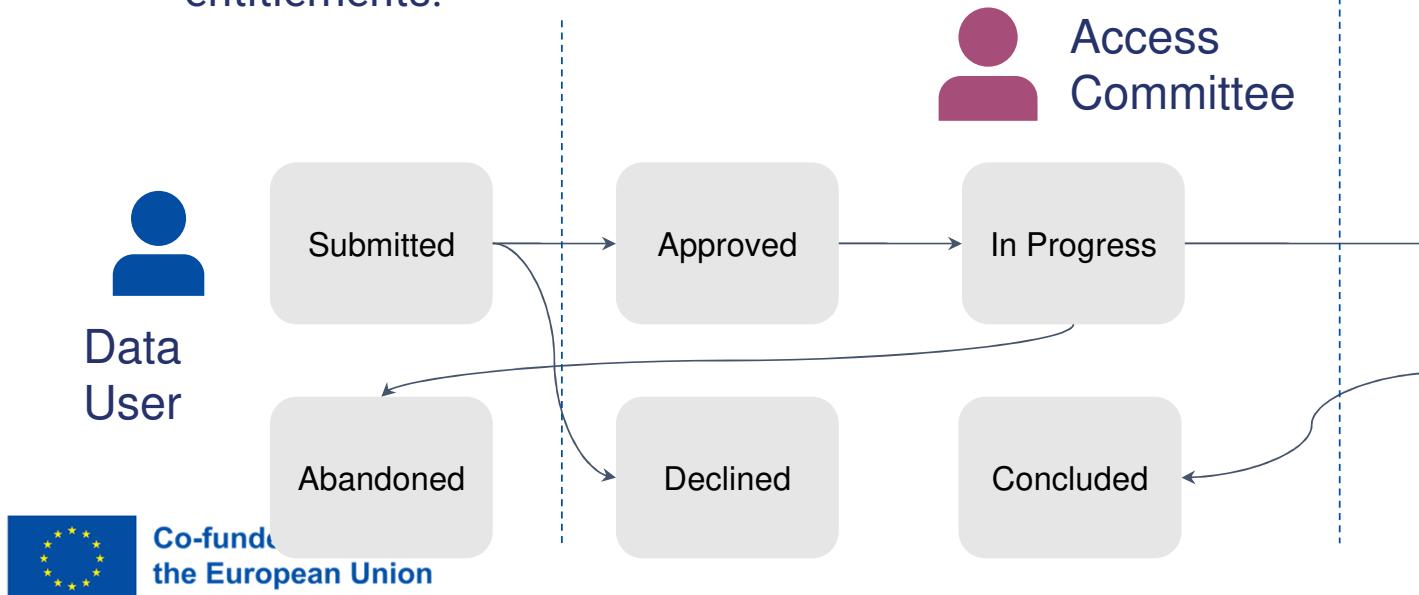


Co-funded by
the European Union



Permit Application

- EUCAIM uses BBMRI-ERIC negotiator.
 - Fine-grain, suitable for multiple stakeholders.
- It relies on LS-AAI for authentication and authorisation.
 - Additional permission information may be added as entitlements.



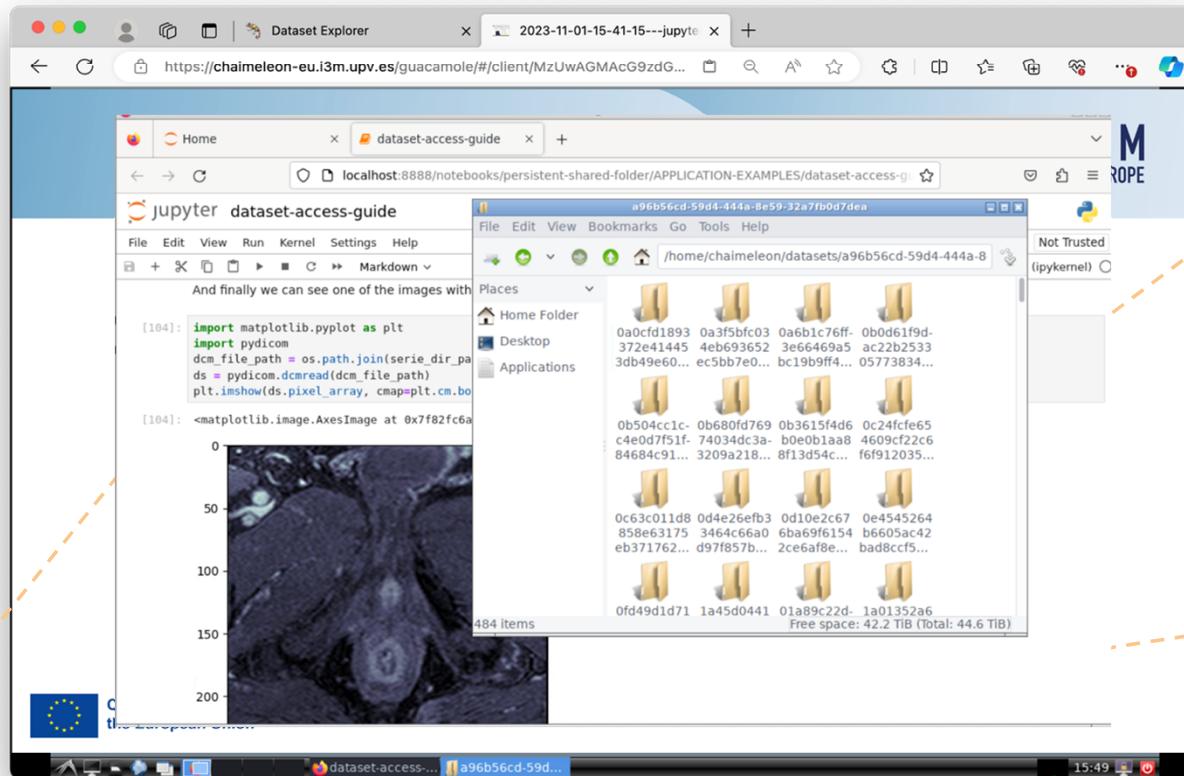
| Dataset Responsible |
|--|
| ☒ Submitted |
| ⌚ Representative Unreachable |
| ✉ Representative Contacted |
| ⟳ Returned for Resubmission |
| 🔍 Checking Availability |
| 📦 Resource Unavailable, Willing to Collect |
| ✗ Resource Unavailable |
| ✓ Resource Available |
| ❗ Access Conditions Indicated |
| ⚡ Access Conditions Met |
| ✗ Resource Not Made Available |

Data use: Virtual Research Environment



A GUI with an
ubuntu
container in a
network-
restricted
environment

Data Analytic
SW libraries



Mounts the
studies of the
datasets as a
POSIX volume.

Link to a batch
queue system
with GPUs and
powerful
resources.

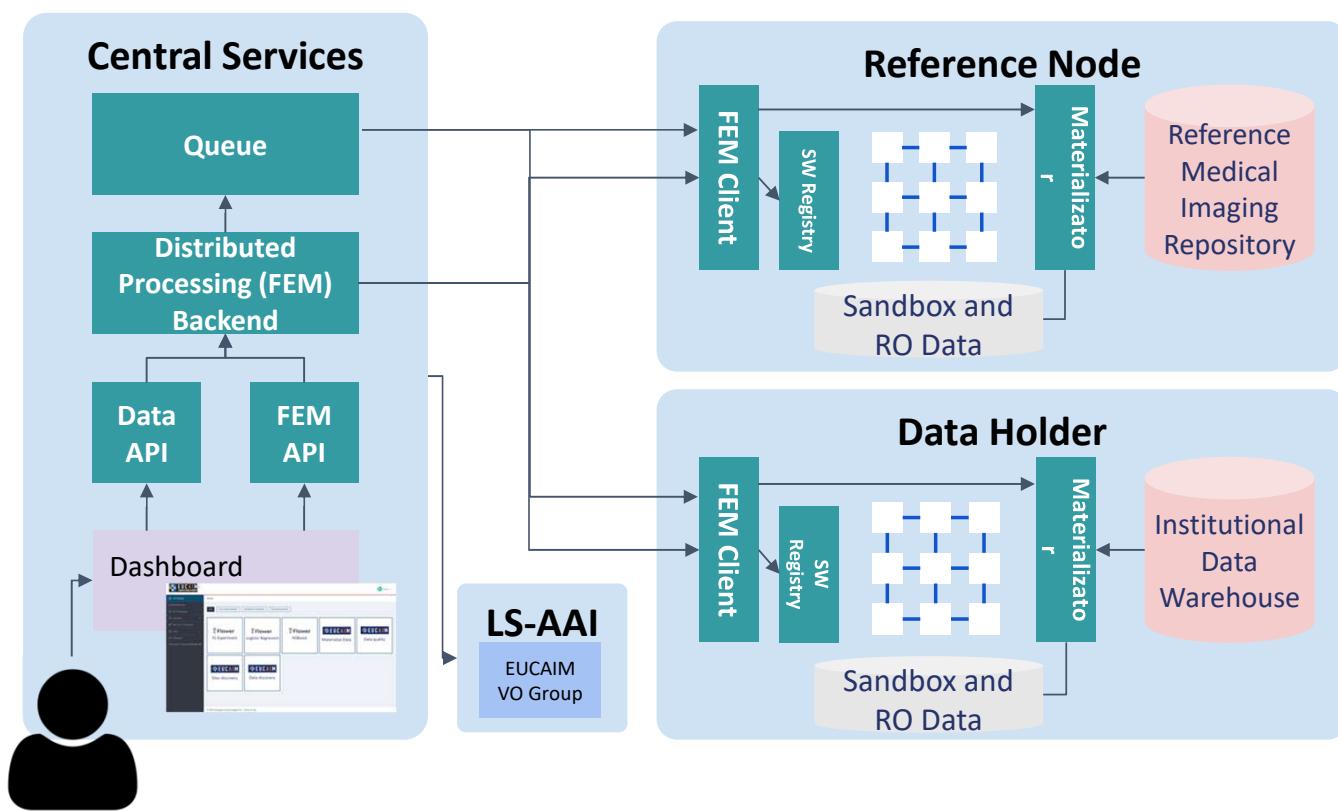
 Data cannot be downloaded due to the network policies and the restricted configuration of the access proxy - accesses are traced and preserved.



Data Use: Federated Processing Environment



- Integration with processing environments have been proven through the demonstrators
- Future integration will focus on:
 - GUI in the Dashboard.
 - Permissions through the User's Library.
 - Integration in the SW Catalogue repository (harbor.eucaim.cancerimage.eu).

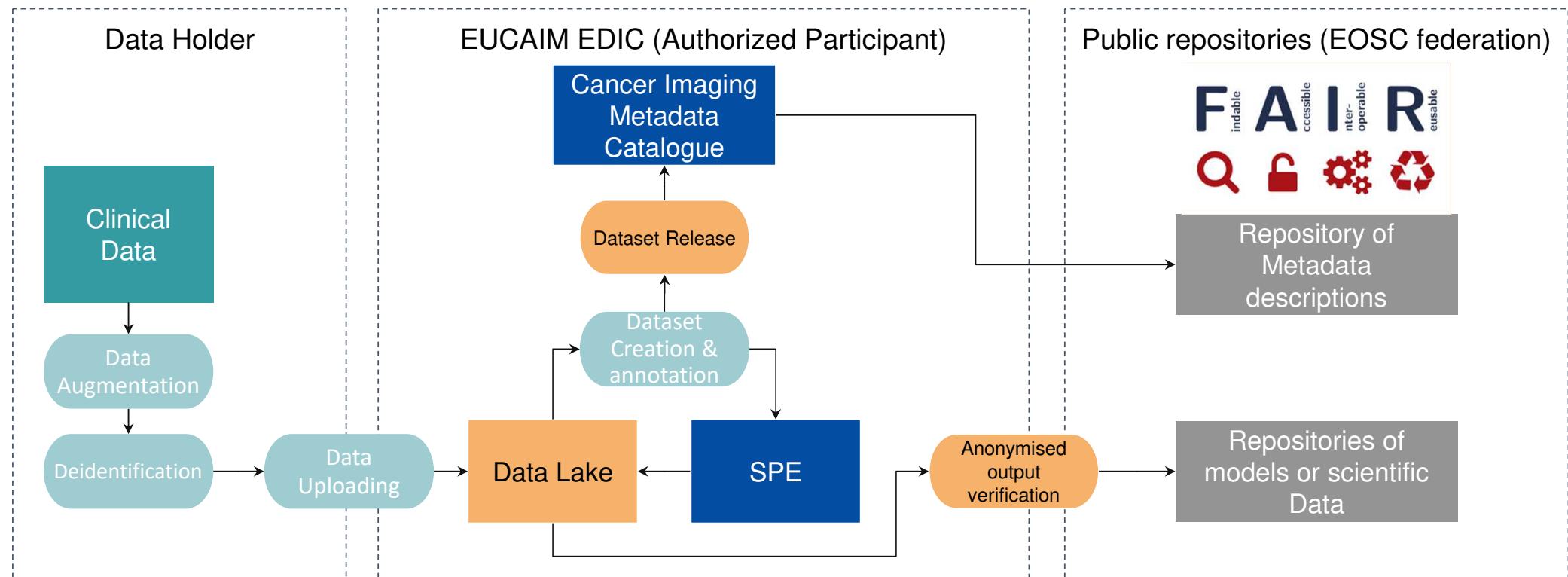


UNIVERSITAT DE
BARCELONA



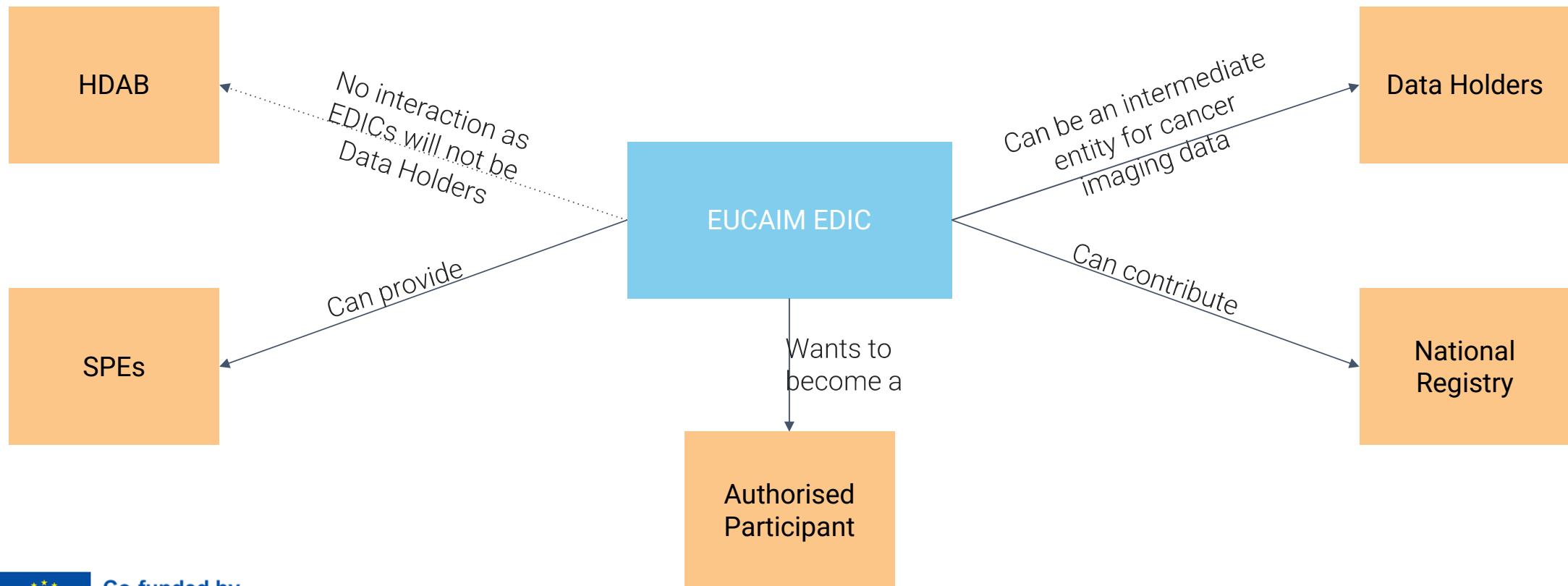
Co-funded by
the European Union

Resource Outputs



Co-funded by
the European Union

EUCAIM and EHDS roles



Co-funded by
the European Union

Gaps and complementarities



| | EHDS | EUCAIM | Comp |
|--------------------------------|--|---|---------------------|
| Architecture | Based on a model of federation of data and service providers. | Based on a model of federation of data and service providers. | same |
| Legal framework | Agreed by countries. | Based on data collection, dataset specific | compl |
| Targeted Data | Mainly sensitive data | Mainly sensitive data | same |
| Scope | Mainly thematic (Health and wellness). | Cancer imaging and associated data | subset |
| Implementation timeline | Member States shall comply with the EHDS by 2029. | Entering production stage | adv. |
| Implementation approach | Hierarchical (country driven). | Mainly bottom-up with top-down guidance. | compl. |
| Data User Journey | Data discovery, Permit application, Data preparation, Data provision, Data use & Results output. | Same for observational studies, simplified data preparation and provisioning in the access to retrospective data. | compl. |
| Implementation force | EHDS construction is compulsory to all the Member States and implementing acts will define specific components (as well as projects). | EUCAIM and some national activities. | align need ed |
| Usage cost model | EHDS preparation of data and services can require a cost, as well as the resources of the SPEs. It is expected that the user will contribute to this cost. | EDIC model combining country support and project based funding. | align need ed |

Functionalities of potential interest for the EHDS



- Many of the components in the platform are aligned with the services requested in the EHDS.

The screenshot shows the EUCAIM Catalogue interface. At the top, there's a search bar and filter options for conditions, topography, body part, geographical coverage, collection type, collection method, order of magnitude, sex, modality, vendor, and image access type. Below the search bar, it says "Dataset series: 45" and "Datasets: 87". The main area displays a table of datasets with columns for Title, Dataset (version) ID, Flags, Author, Project ID, Created on, Studies, and Subjects. One entry is highlighted: "Lung Cancer Training/Validation dataset (October 2024)" by Ana Jerez-Pastor, CHAMELEON, created on 01/10/2024, with 1088 studies and 1088 subjects.

Registries

The screenshot shows the EUCAIM Federated Data Explorer interface. It features a search bar for "Modality: Computed Tomography" and "Modality: Magnetic Resonance Imaging". Below the search bar, it says "Provider: 2 / 2" and "Studies: 25490" and "Subjects: 23235". The main area displays a table of datasets with columns for Patient, Clinical Parameters, and Image Parameters. One entry is highlighted: "CHAMELEON Lung + breast" by CHAMELEON, created on 01/10/2024, with 1088 studies and 1088 subjects.

Data Discovery

The screenshot shows the EUCAIM Negotiations interface. It displays a list of negotiations with columns for Negotiation ID, Created on, Author, and Status. One negotiation is highlighted: "CHAMELEON Lung + breast" by Ignacio Barquer Expert, created on 20/09/2023 at 11:23. Another negotiation is shown: "CHAMELEON Colon DO NOT APPROVE" by Ignacio Barquer Expert, created on 20/09/2023 at 13:46.

Data Access

The screenshot shows the EUCAIM Secure Processing Environments interface. It includes a "jupyter notebook" tab showing a 3D CT scan of a lung, a "dataset access" tab showing a 2D histology image, and a "dataset access" tab showing a 3D MRI scan. There are also tabs for "Diagnose", "Data", and "Work Progress".

Secure Processing Environments



Co-funded by
the European Union