



# ELIXIR Human Data Use Case

*Mikael Borg, ELIXIR Sweden*



ELIXIR-EXCELERATE is funded by the European Commission within the Research Infrastructures programme of Horizon 2020, grant agreement number 676559.

*[www.elixir-europe.org/excelerate](http://www.elixir-europe.org/excelerate)*

# What is the EGA?

The EGA is an Archive for human genomic data that should have controlled access because it allows identifying individuals

- ✓ Data is provided by research centers and health care institutions
- ✓ Access is controlled by Data Access Committees
- ✓ Data requesters are researchers from other research or health care institutions

*Our goal is to foster responsible data reuse for the good of science and all humankind*

<https://www.ebi.ac.uk/ega>  
<https://ega-archive.org>

The image displays two overlapping screenshots of the European Genome-phenome Archive (EGA) website. The top screenshot shows the main landing page with a navigation bar (HOME, BROWSE, REQUEST, SUBMISSION), a search bar, and a 'What is in the EGA?' section featuring a bar chart of studies by disease type. The bottom screenshot shows a detailed view of the 'European Genome-phenome Archive' page, including a search bar, a list of datasets, and a 'Help' section with links to FAQs and contact information.

**What is in the EGA?**  
Studies in the EGA by disease  
Click on a column to view category subgroups

Disease Type	Number of Studies
Cancer	582
Cardiovascular	92
Infectious	22
Inflammatory	99
Neurological	13
Other	133

If applicable, a study may be included in more than one category

**Latest studies**  
Efficacy of JAK/STAT pathway inhibition in murine xenograft models of early T-cell precursor (ETP) acute lymphoblastic leukemia. - March 12, 2015

**European Genome-phenome Archive**  
Services | Research | Training | About Us

Search: [All] [Search]

Examples: EGA0000000001, Cancer

**EGA home** | About | Studies | Datasets | Data access committees | Data providers | Submit to EGA | Contact Us

**Help**

- Users FAQ
- Submitters FAQ
- Using your EGA account
- Contact Us
- EGA mailing list

**Studies**  
Studies are experimental investigations of a particular phenomenon or trait.  
[Browse all studies](#)

**Learn about the EGA**

- Introduction to the EGA
- How to obtain an account with the EGA
- Using your EGA account

**Datasets**  
The EGA archives a large number of datasets, the access to which is controlled by a Data Access Committee (DAC).  
[Browse all datasets](#)  
[Browse all control datasets](#)

**Data Access Committees**  
Providers may be involved in study creation, submission and designation of Data Access Committees (DACs).

**Navigation**

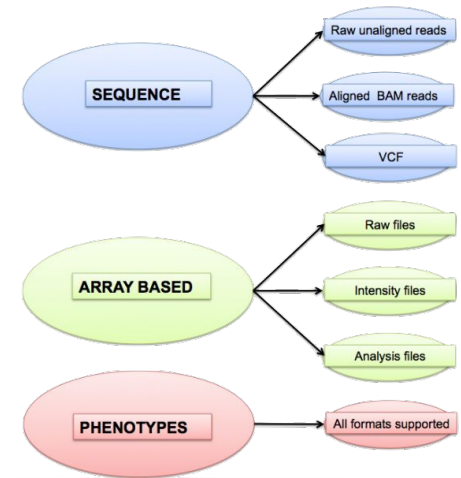
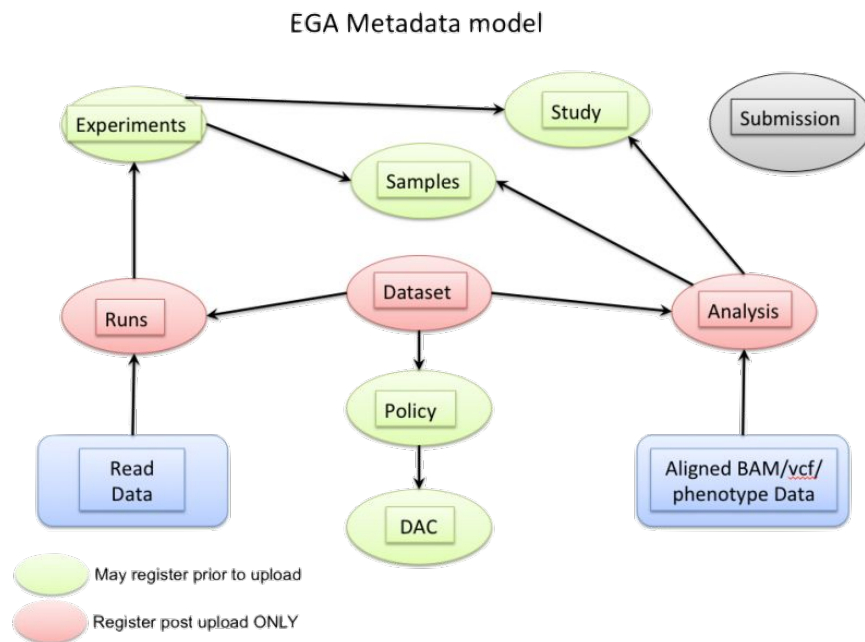
- Login
- Request new password

**European genome-phenome archive**

**elixir**

# EGA: European Genome-phenome Archive

- Repository that promotes the distribution and sharing of genetic and phenotypic data consented for specific approved uses but not fully open, public distribution.
- All types of sequence and genotype experiments, including case-control, population, and family studies.
- Study & Metadata searchable
- Shares parts of the Metadata model with ENA

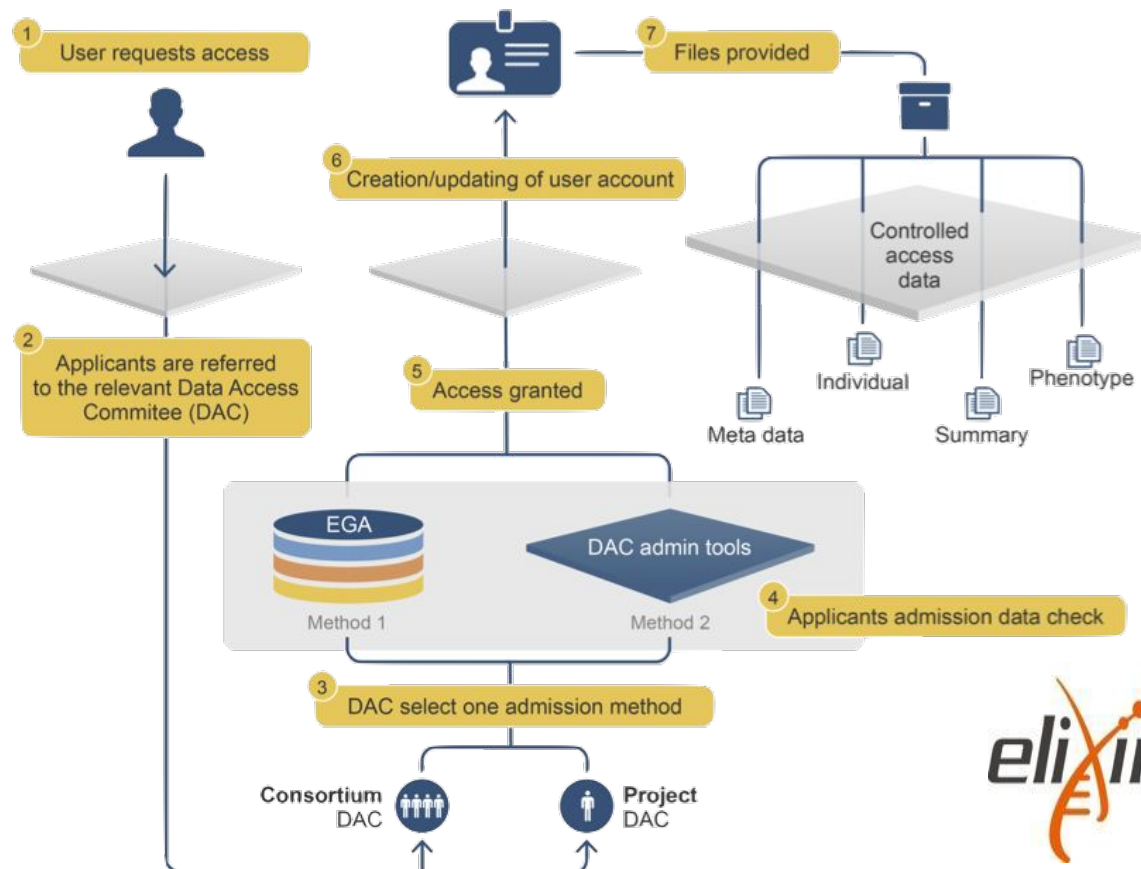


European  
genome-phenome  
archive

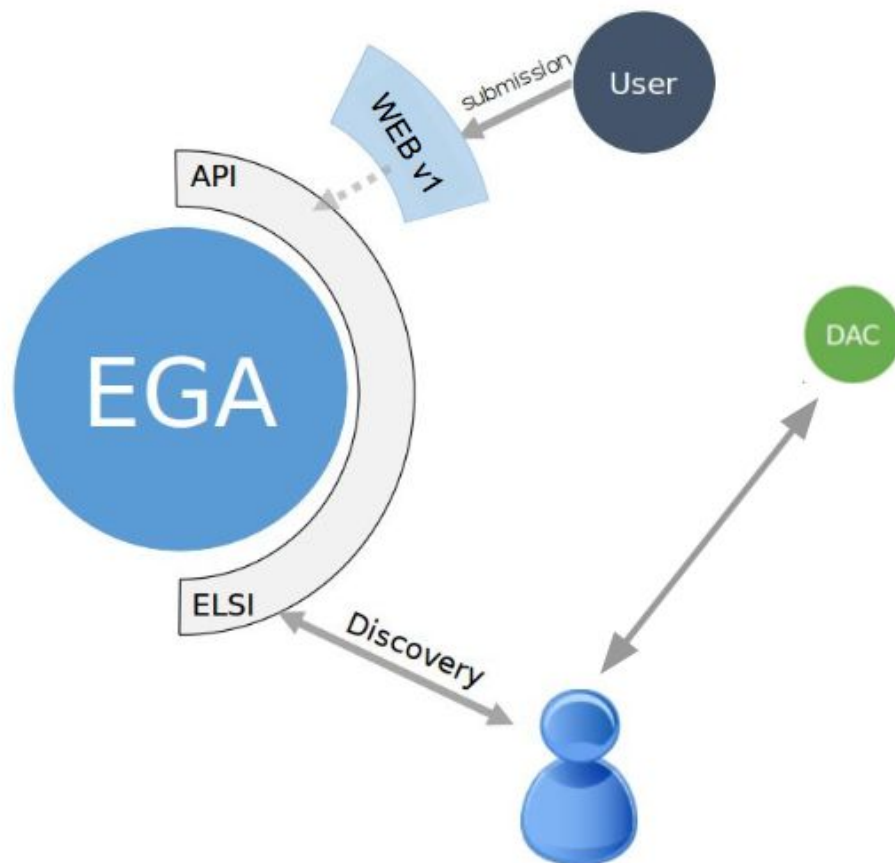


# EGA

- Data Access Agreement
  - Defined by the dataset owner
- Data Access Committee – DAC
  - Decided by the dataset owner



# Pre-EXCELERATE EGA

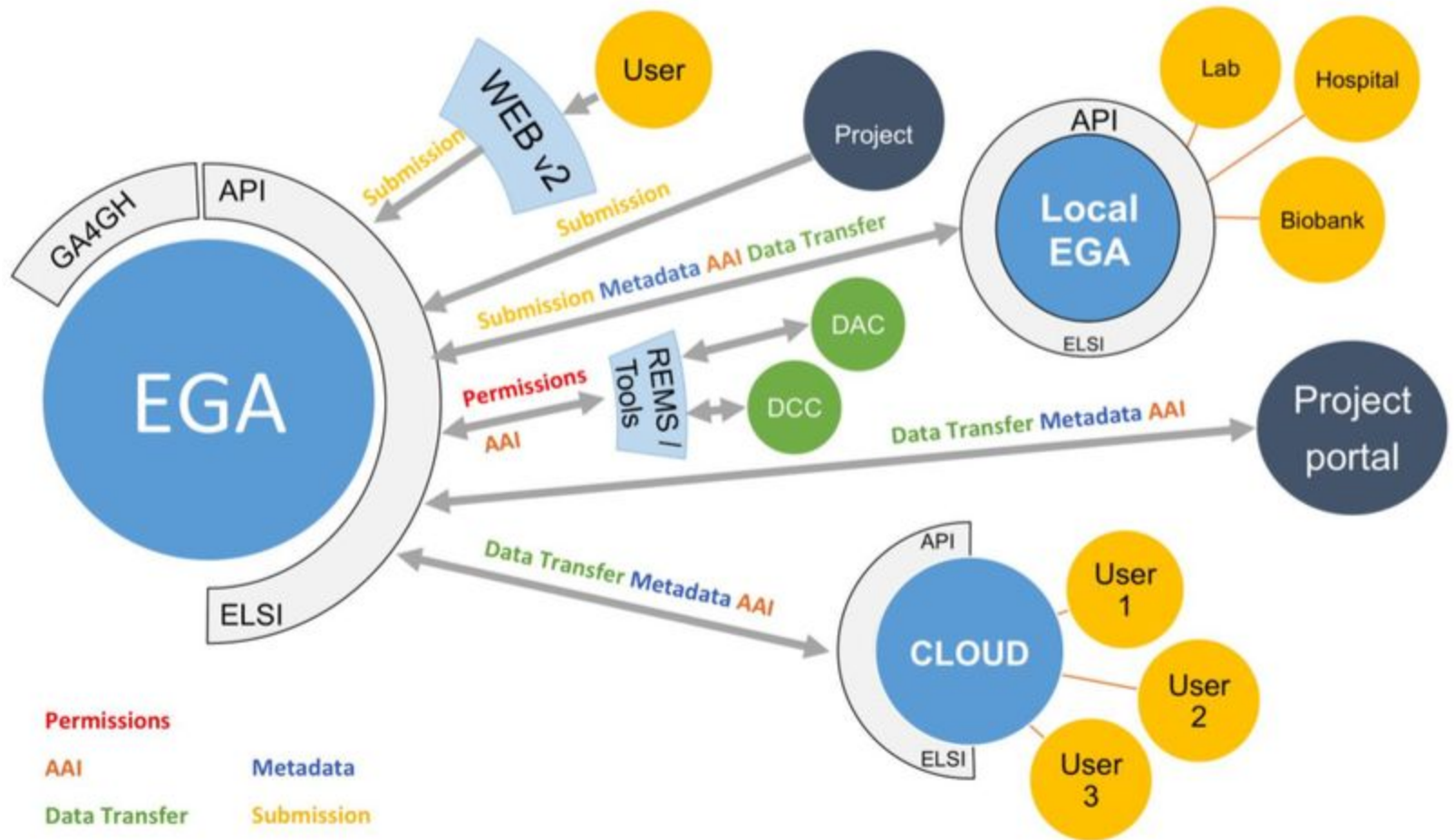


EGA future is

- Distributed
- Cloud based
- Accessible



# Post-EXCELERATE EGA

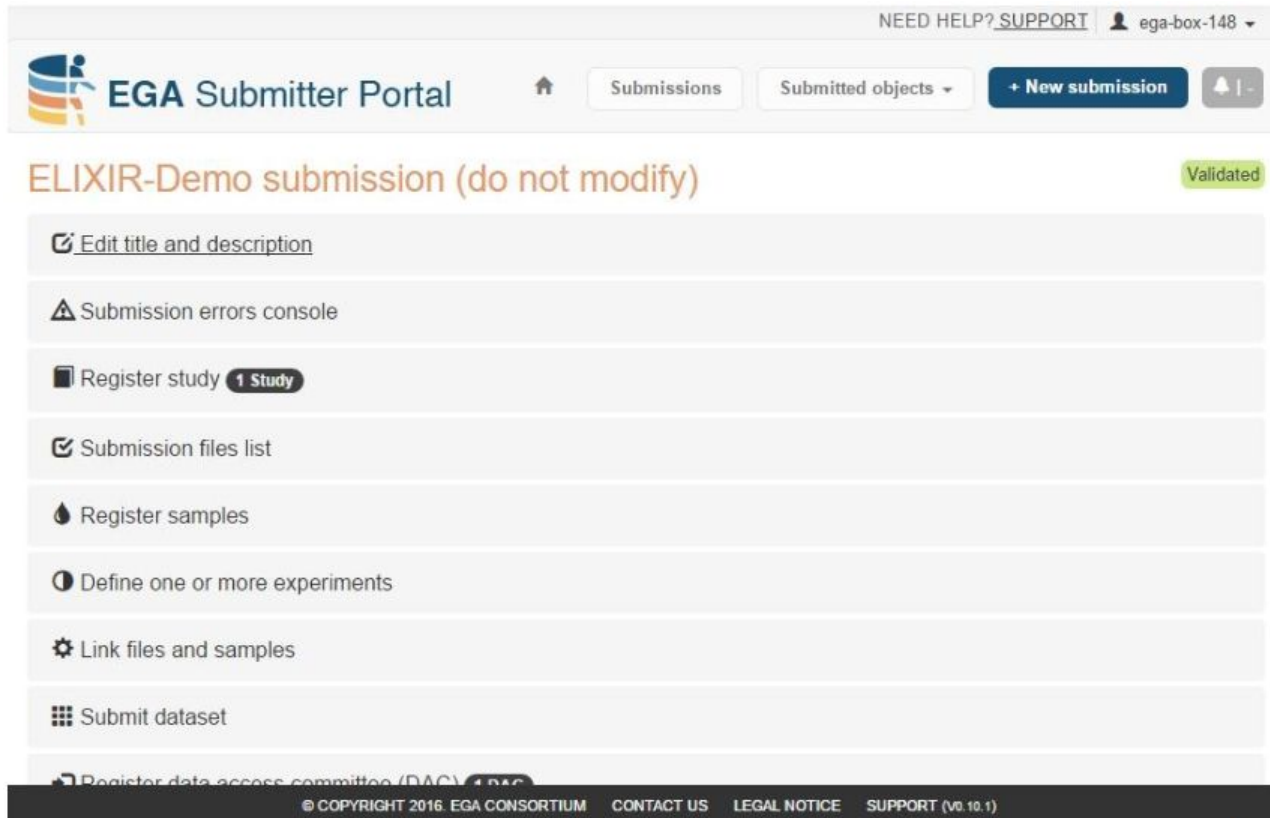


# WP9 Overview

- Task 9.1
  - Support for large-scale submissions of -omic data and sample metadata to the EGA
  - Portable submission toolkit (“EGA-in-a-box”, “Local EGA”)
- Task 9.2
  - Support secure integration of EGA data to downstream project client websites
  - Access management workflow support
  - ELIXIR and EGA access integration
- Task 9.3
  - Large-scale data mirroring support
  - EGA data access authorization integration
  - Data access APIs



# Submission UI



The screenshot shows the EGA Submitter Portal interface. At the top, there is a header with the EGA logo, the text "EGA Submitter Portal", and navigation links for "Submissions" and "Submitted objects". A "New submission" button is also present. The main content area displays "ELIXIR-Demo submission (do not modify)" with a "Validated" status. Below this, a list of actions is provided: "Edit title and description", "Submission errors console", "Register study" (with a "1 Study" badge), "Submission files list", "Register samples", "Define one or more experiments", "Link files and samples", "Submit dataset", and "Register data access committee (DAC)". The footer contains copyright information and links for "CONTACT US", "LEGAL NOTICE", and "SUPPORT (v0.10.1)".

UI on top of the  
Submission API

Available at  
<https://ega-archive.org/submitterportal-test/>

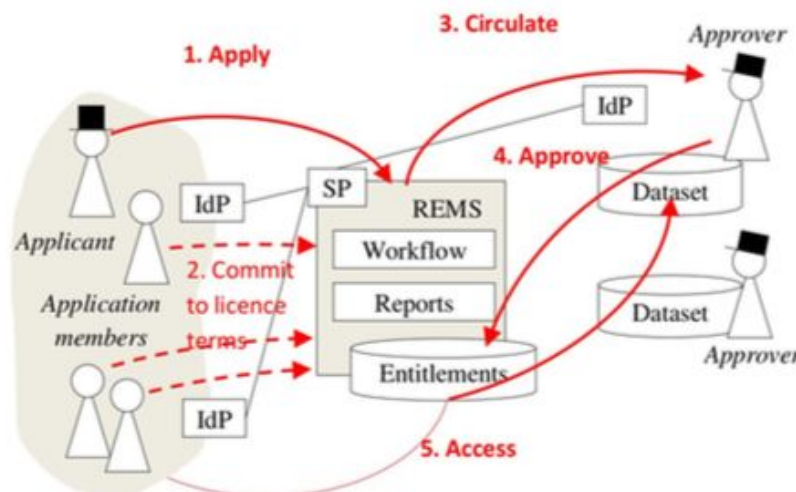
User/pwd:  
ega-box-148 / ELIXIR-Demo



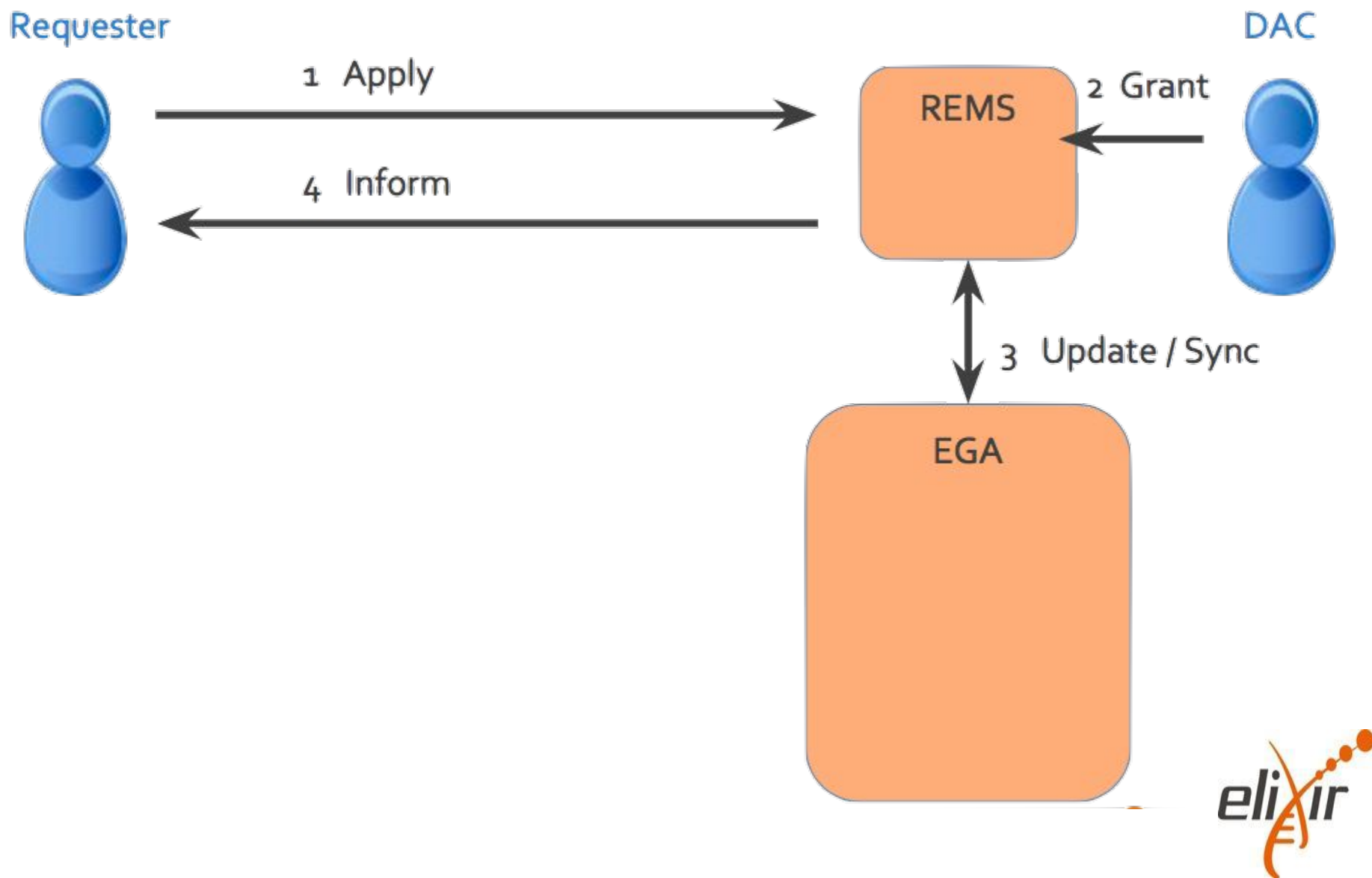


# Access management workflow support (9.2.2)

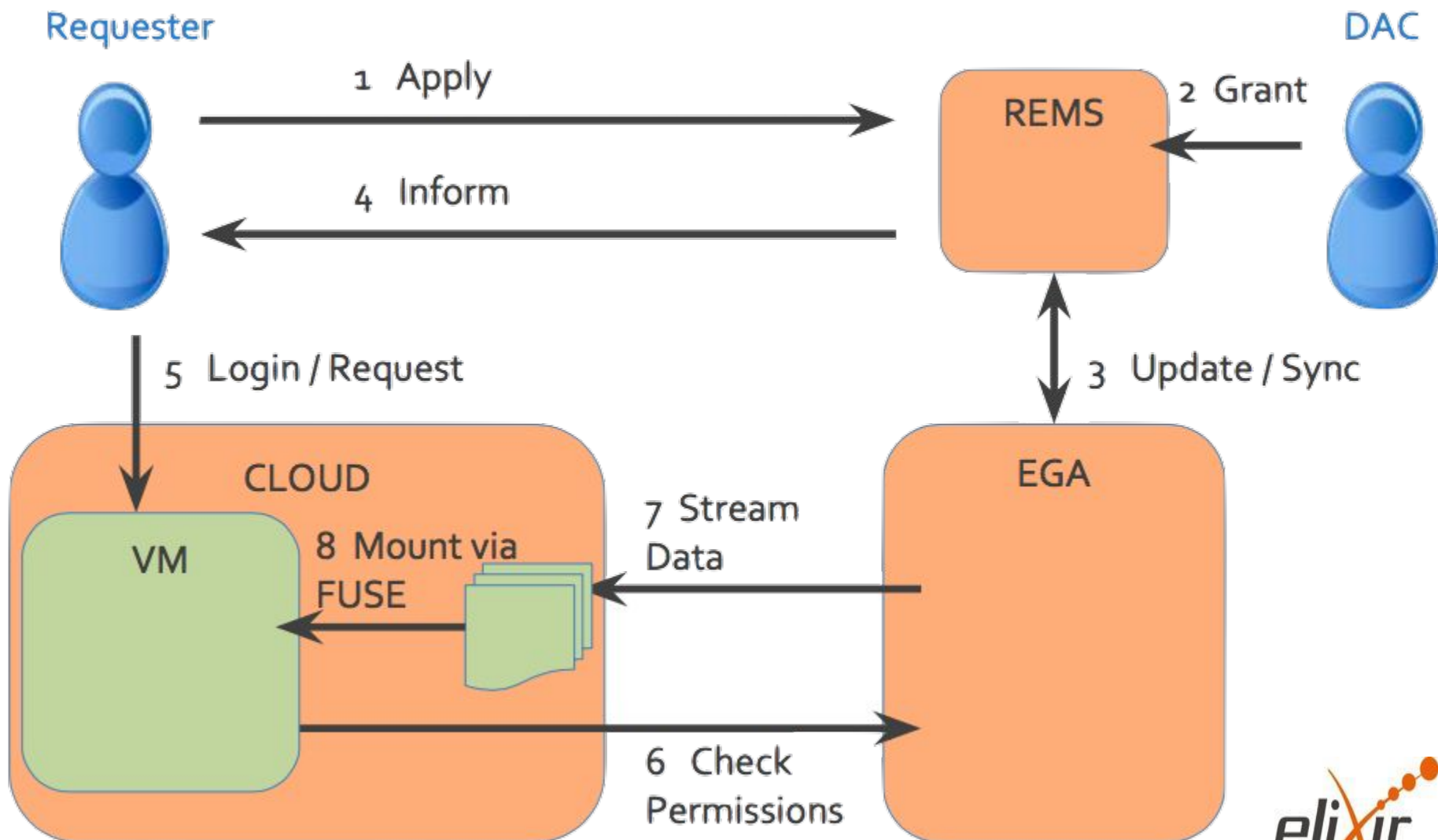
- Resource Entitlement Management System (REMS)
  - Electronic tool for the management of access rights to research data
  - Supports federated identity authentication for both parties
  - Developed at the CSC Finland
- Apply for an access to a resource
  - Login using their home organisation's user ID and password
  - Fill in an electronic data access application
  - Agree to the terms of use for this dataset
- Circulate data access application for approval
  - Provides reports on applications and approved access rights



## Access management workflow support (9.2.2)

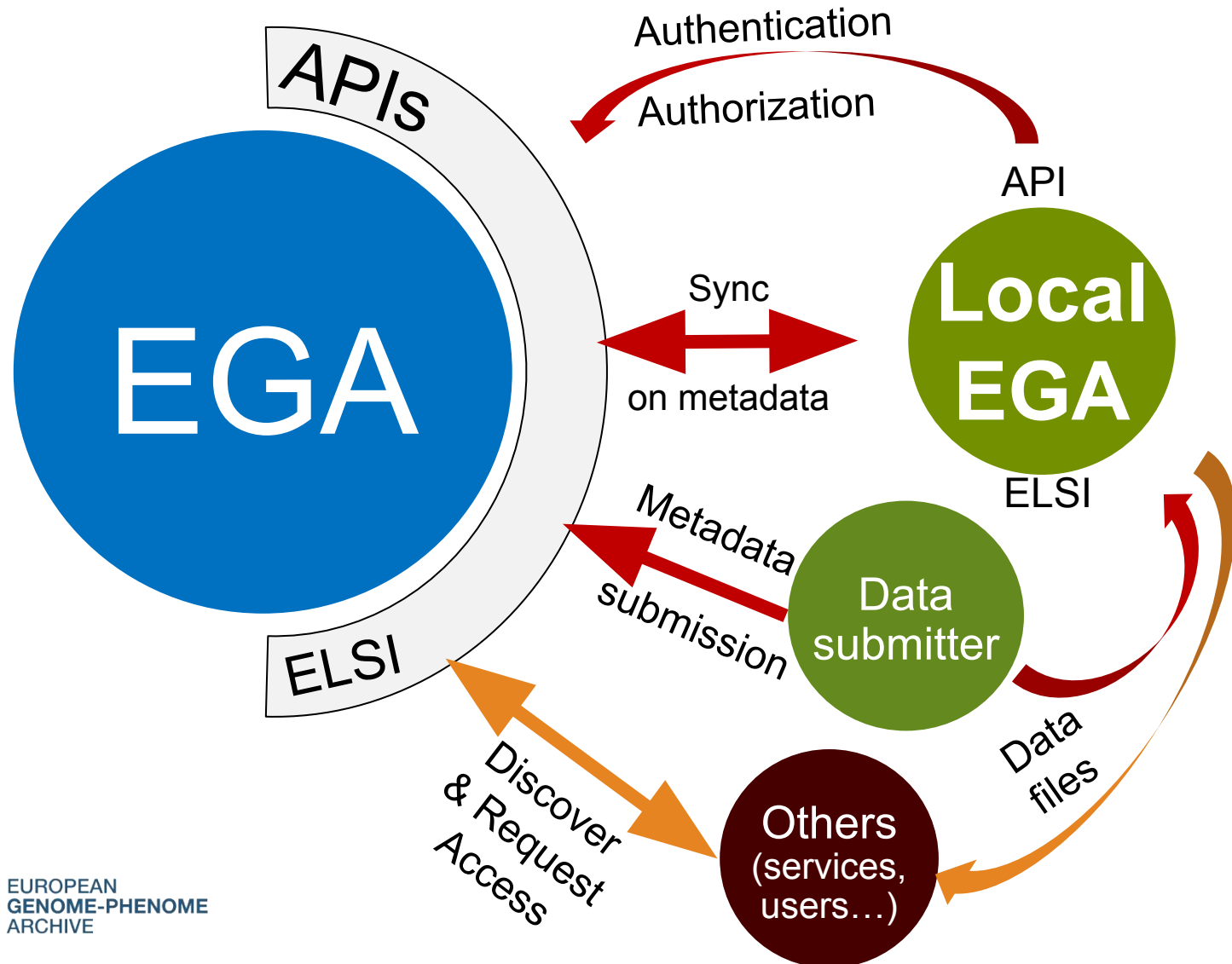


# Access management workflow support (9.3.2)



Secure cloud access demo movie: <https://www.youtube.com/watch?v=gwkZzz1e8bY>

## Task 9.1.2: Local EGA



# Local EGA Goals

- Establish a “portable” core code base, that supports
  1. **Submitter Registration** at CEGA, specifying LEGA instance to store data
  2. **Dataset Ingestion and Archiving**
  3. **Requester Registration & Dataset Access**

<https://github.com/NBISweden/LocalEGA>

- *Optional processes*
  1. On-the-fly decryption data access in secure cloud instances (9.3.2)
  2. Mirroring/caching of CEGA hosted datasets at a LEGA instance (9.3.1)
  3. *Sync of metadata between CEGA and LEGA instance*

- **Hardware agnostic**

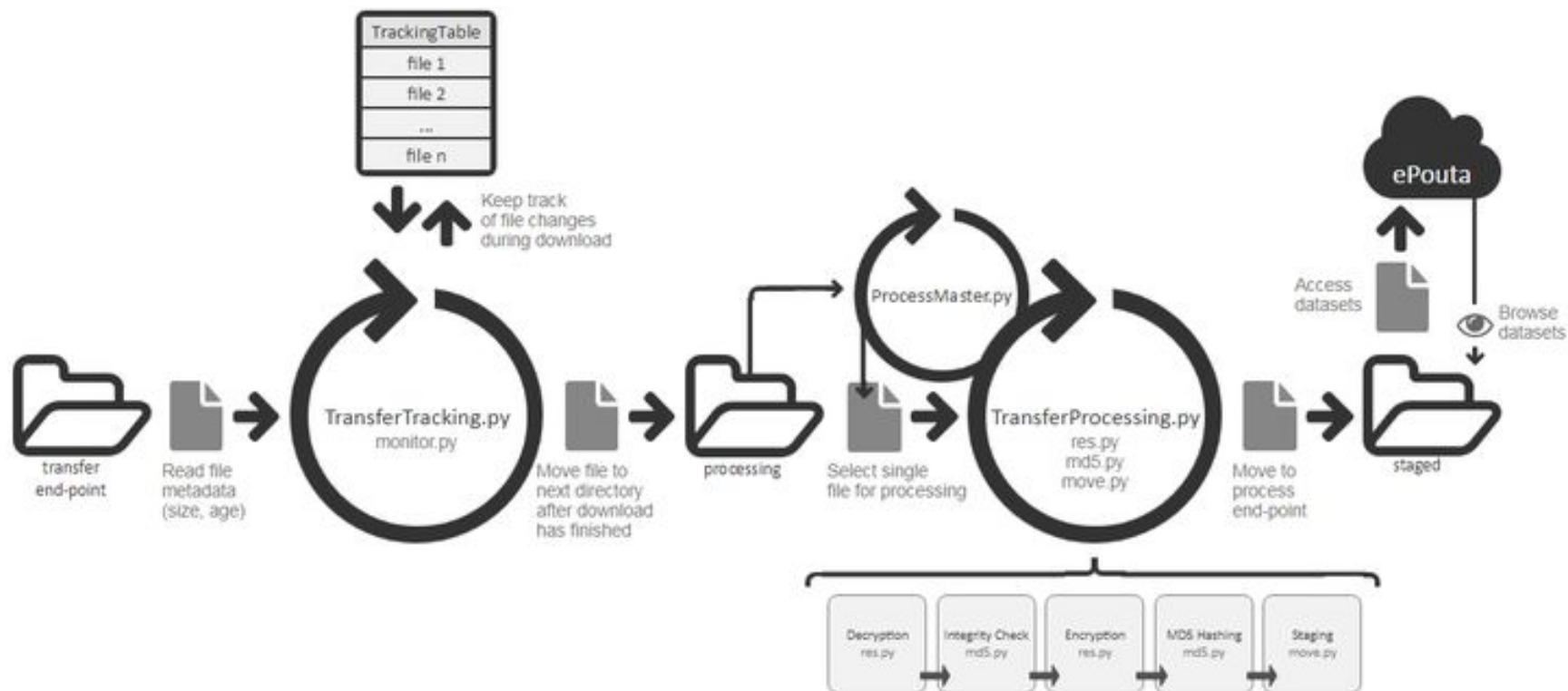
- Nordic partners working on establishing Local EGA instances

- Coordinated in the Nordic [Tryggve project](#)

- *Infrastructure for collaboration on sensitive personal data for biomedical research*



# Data mirroring (9.3.1)



- Series of Python scripts in Luigi workflows

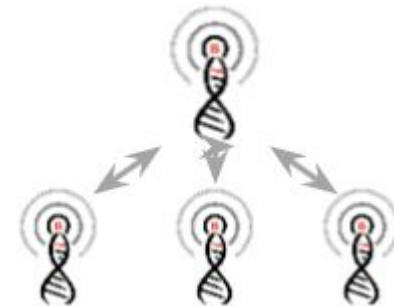
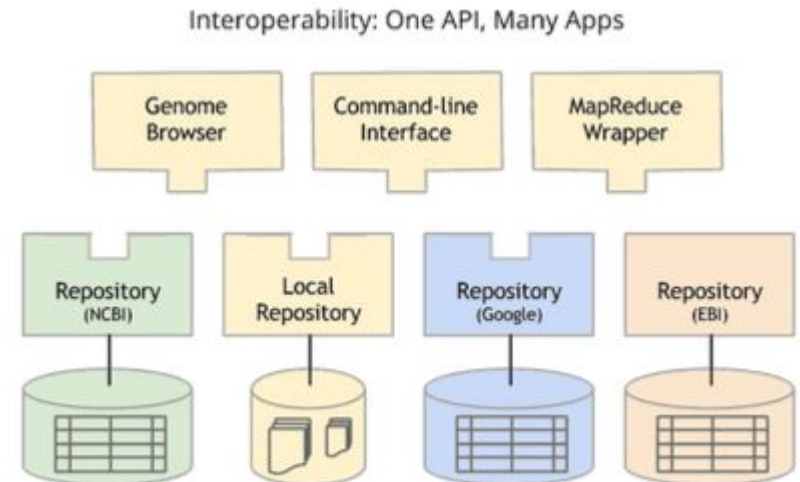
[github.com/CSCfi/lega-mirroring](https://github.com/CSCfi/lega-mirroring)



# Data access APIs (9.3.3)



- Data discovery
  - Beacon API
  - Matchmaker API
- Data delivery
  - Reads and variants API
  - Streaming API
- Secure storage at rest
  - File formats encryption container API
- Security working group
  - Software security task team - AAI recommendations



# Acknowledgements

## **WP9 Partners**

ELIXIR Finland

ELIXIR Sweden

ELIXIR Norway

ELIXIR Denmark

ELIXIR Estonia

ELIXIR Netherlands

EMBL EBI

ELIXIR Spain

## **Other nodes contributors**

ELIXIR Switzerland

ELIXIR Belgium

ELIXIR Luxembourg

ELIXIR Slovenia

ELIXIR Czech Republic

ELIXIR France

## **The ELIXIR Hub**





[www.elixir-europe.org](http://www.elixir-europe.org)



@ELIXIREurope



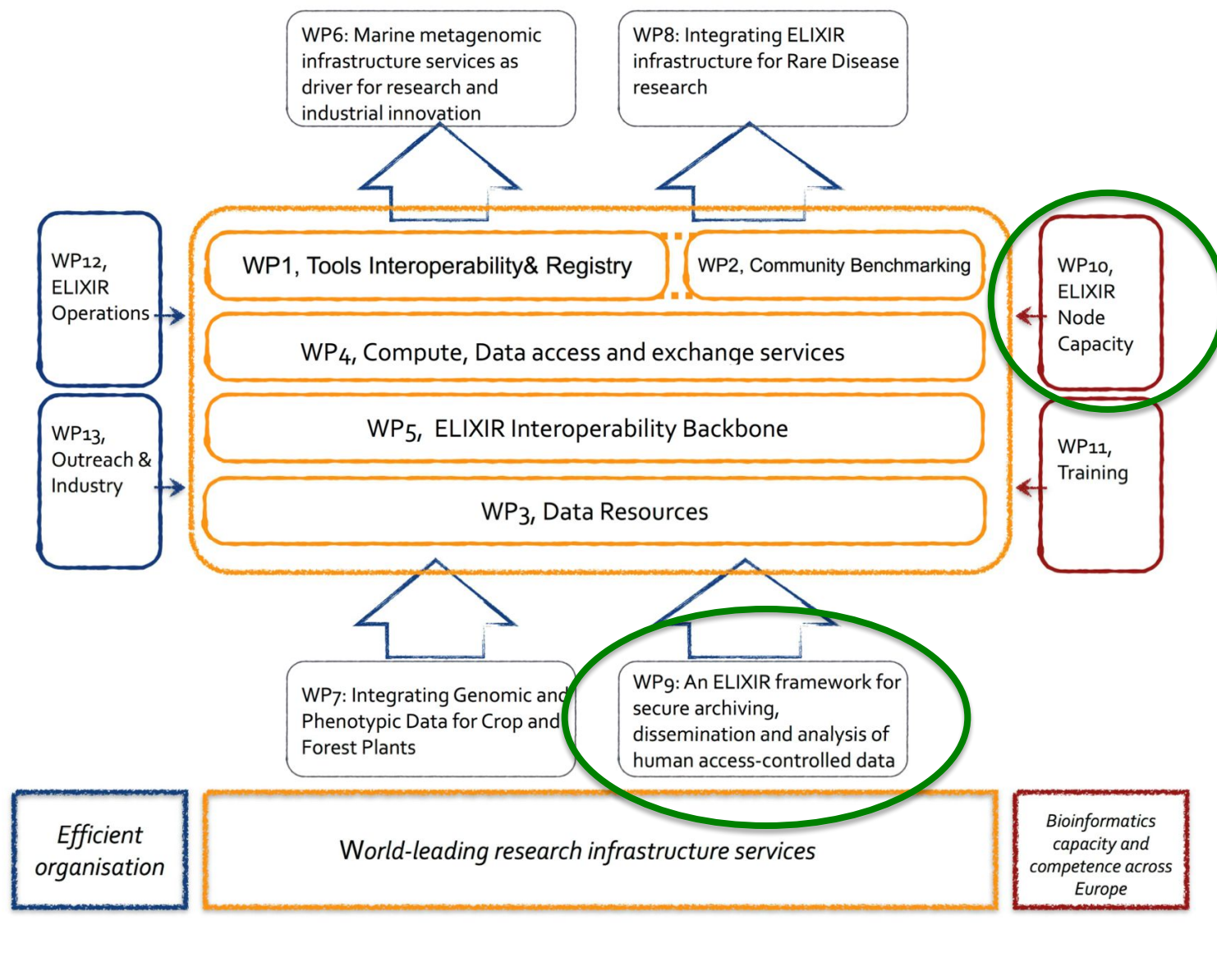
/company/elixir-europe



**Alt slides**

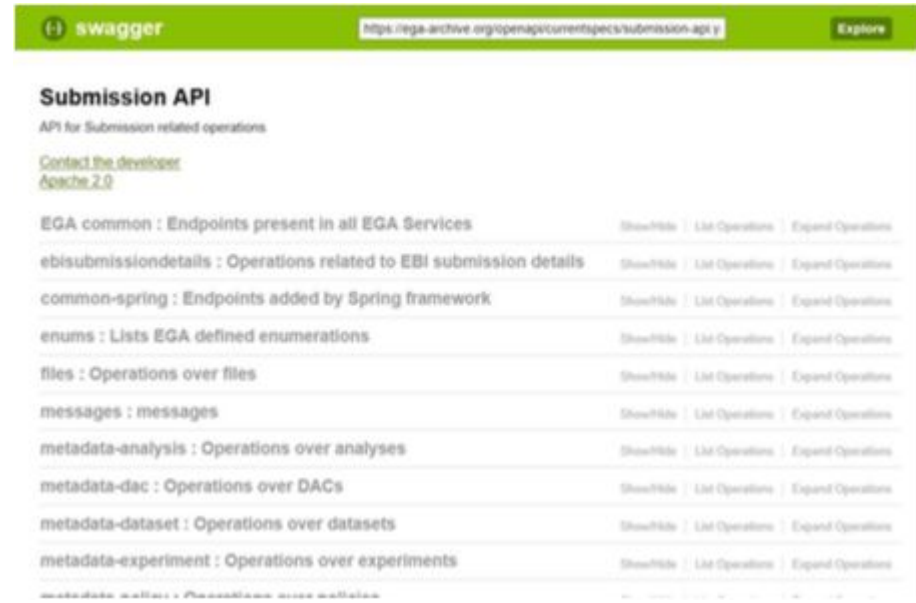


# ELIXIR - Excelerate



# Submission API and Submission UI (9.2.1)

- Creating, editing and validating submissions prior to submission
- EGA submission
- OpenAPI documentation available at
  - <https://ega-archive.org/openapi/currentspecs/submission-api>
- Supports JSON and XML

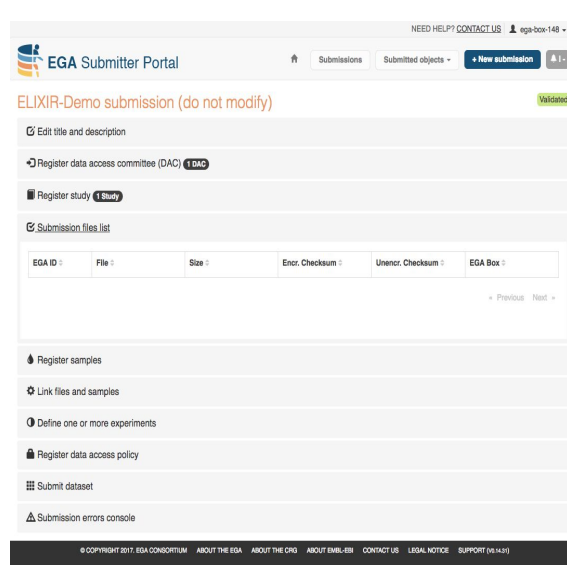




# User registration & Data Archiving

## Central EGA

## Local EGA instance



Create INBOX for user X

INBOX created for user X

Archive FILES for user X

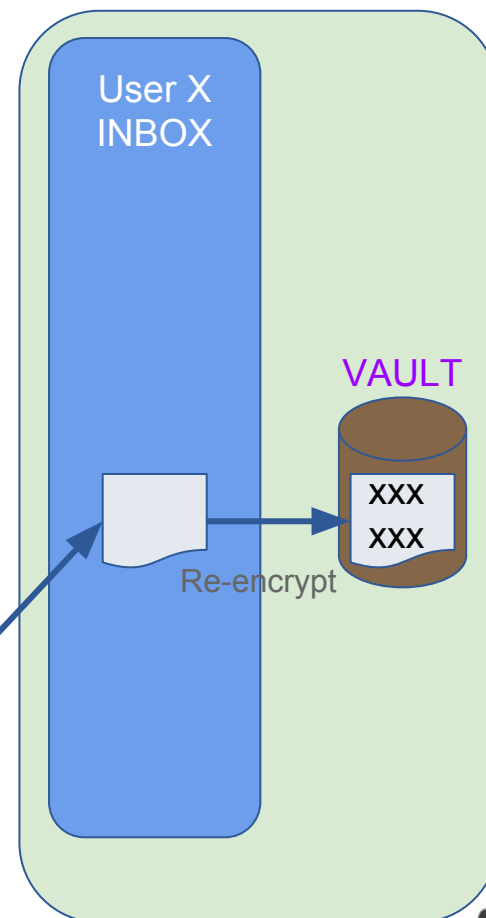
FILES archived for user X  
or ERROR

Make  
submission

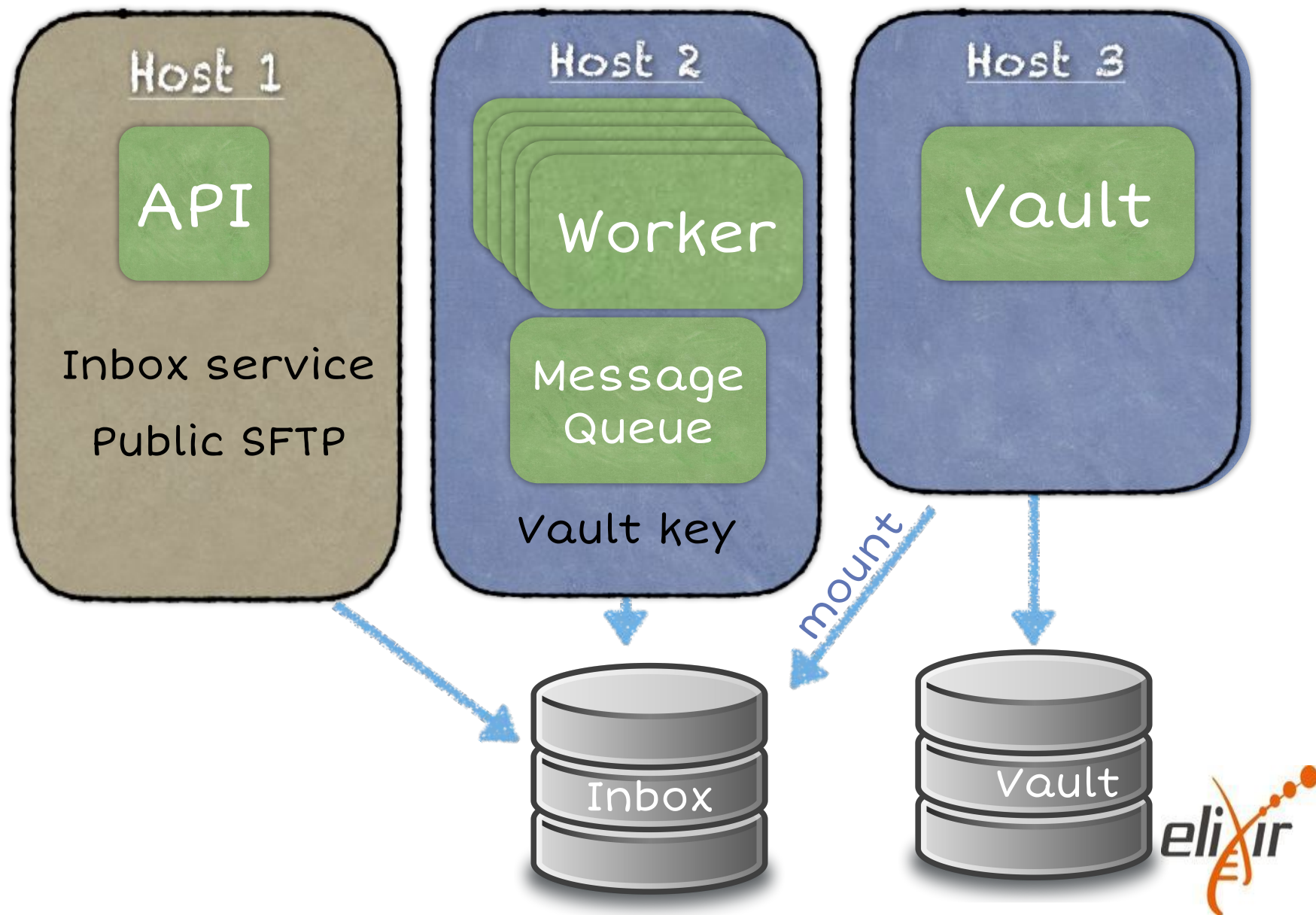
Register  
Submitter



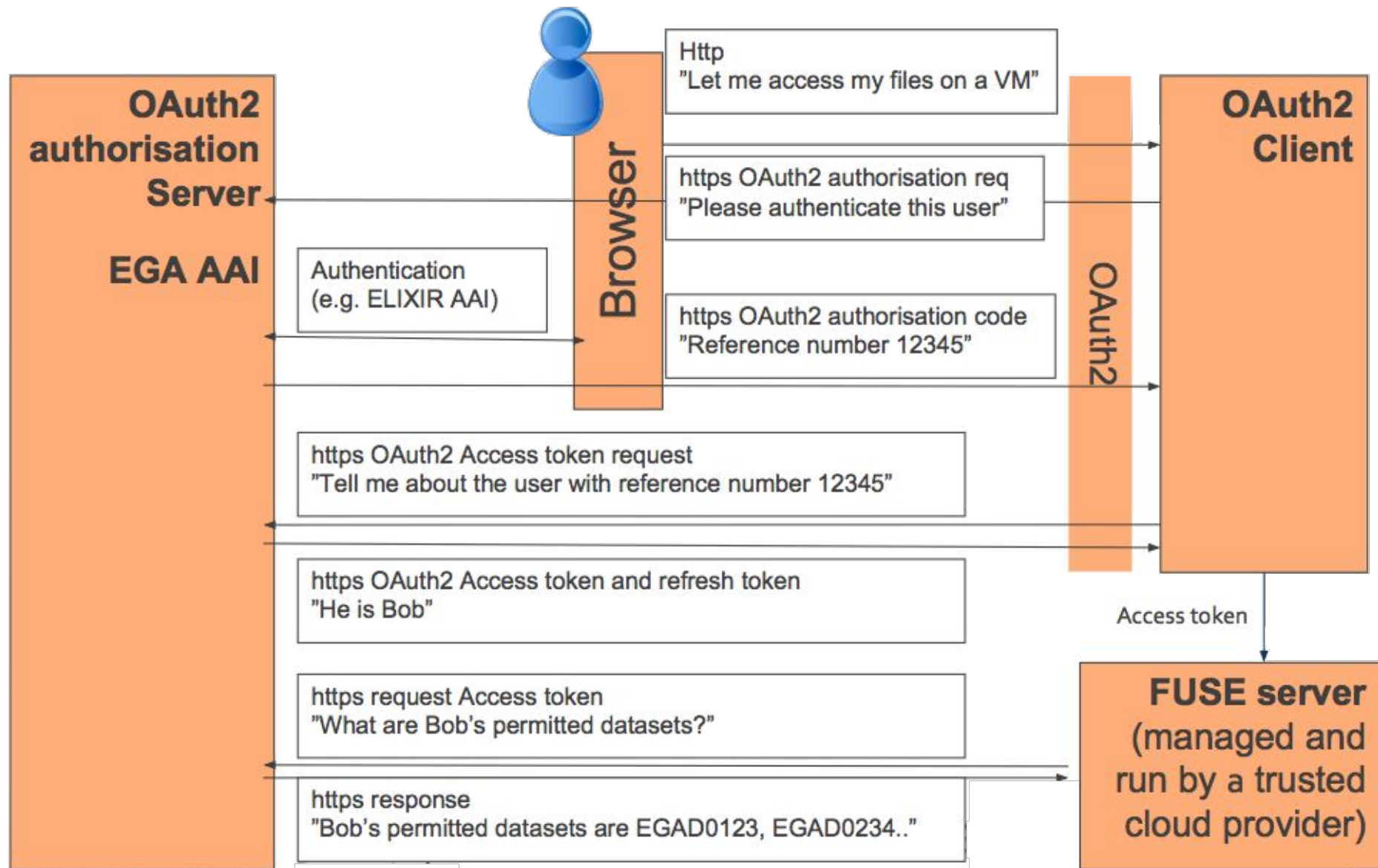
Submitter



# Local EGA Architecture



## ELIXIR and EGA access integration (9.2.3)



# Data access API

