Intelligent Data Ecosystem for Rare Cancers

Principal Investigator: Annalisa Trama
Collaborator: Paolo Baili
Fondazione IRCCS Istituto nazionale Tumori, Milan

IDEA4RC has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement no 101057048

What is it about?

- IDEA4RC is a funded Horizon Europe project led by Fondazione IRCCS Istituto Nazionale dei Tumori (Milan, Italy)

- The project will stretch over the course of four years (September 2022 - September 2026)

- Its objective is

  building an intelligent ecosystem to improve the governance, the sharing, and the re-use of health data for rare cancers

IDEA4RC

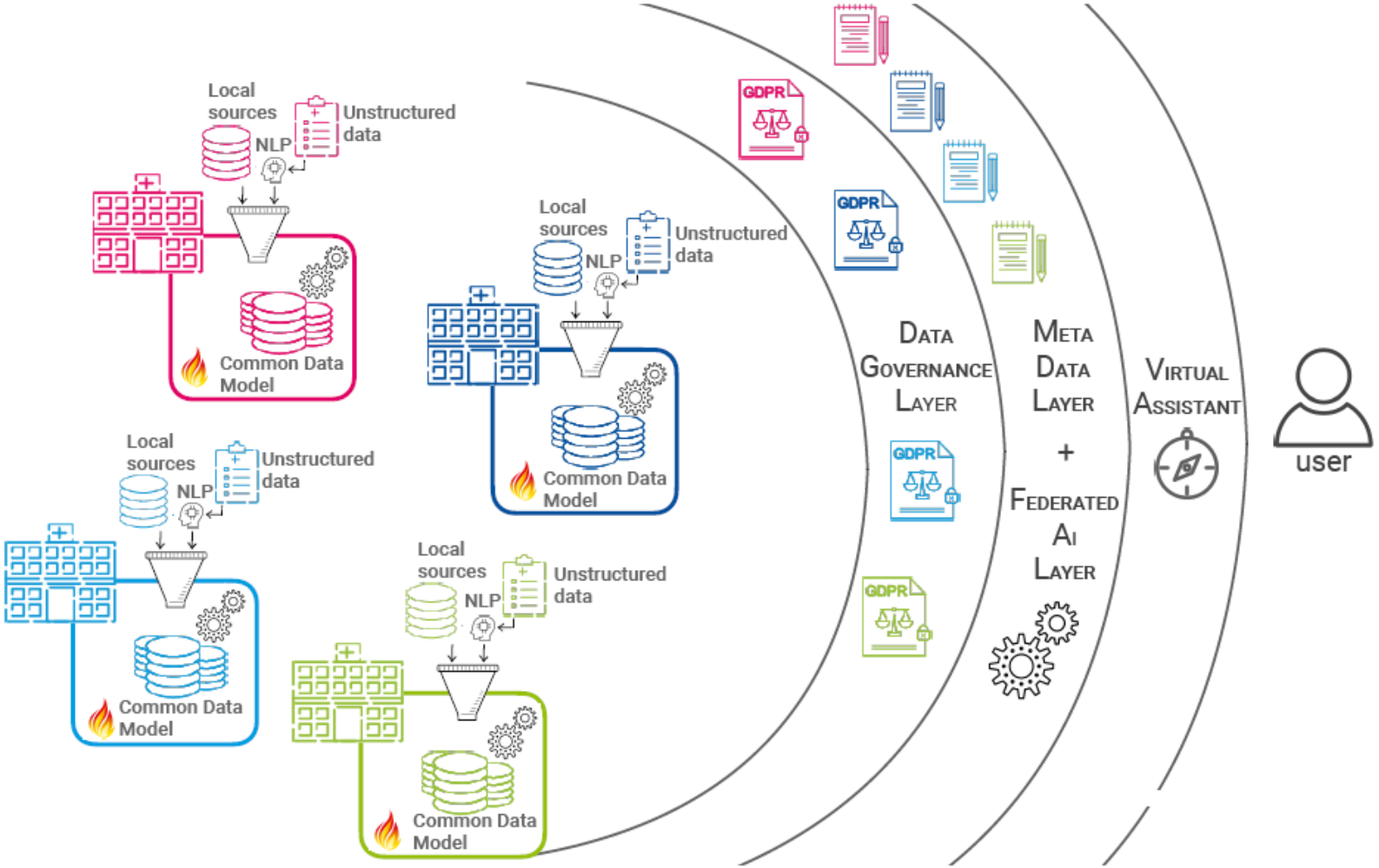
Why do we need such an ecosystem?

- Every year in Europe 650’000 people receive a rare cancer diagnosis. Taken together they represent nearly 25% of all cancer diagnoses in the continent.
- Analysing large and diverse datasets collected by different clinical centers and countries would greatly advance the knowledge on rare cancers.
- Current hurdles include:
  - lack of interoperability
  - compliance with EU data protection requirements when sharing health data.

The principles

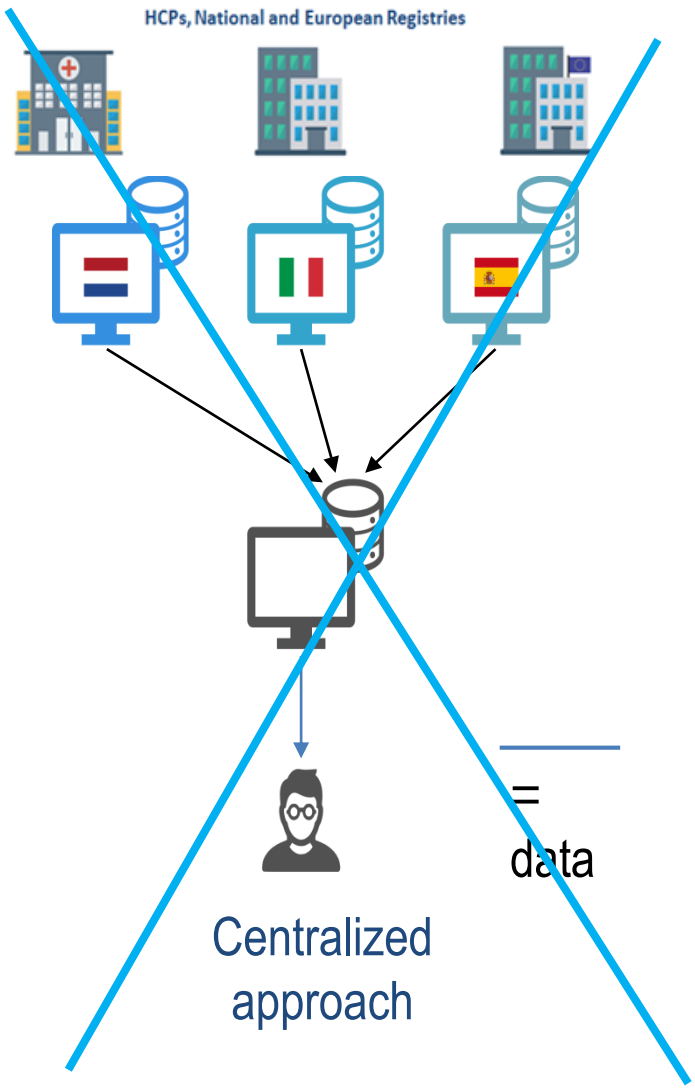
- IDEA4RC plans to develop a new IT infrastructure that implements data protection and privacy by design and by default required by EU regulations and complies with the FAIR principles of scientific data management (Findability, Accessibility, Interoperability, and Reusability).
- The ecosystem will be tested on pilot projects involving 11 centers of excellence of the EURACAN network (EUropean reference network for Rare Adult solid CANcers) partners of the project.

The ecosystem

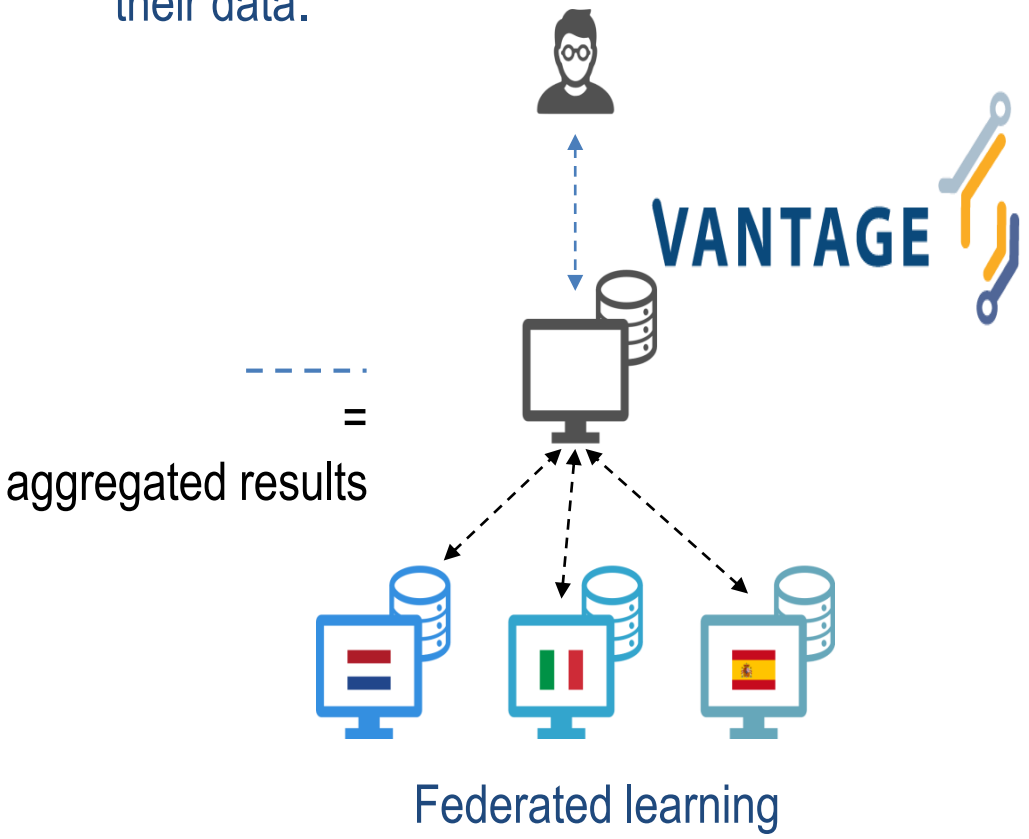


Secure processing environment

Federated approach



Performing an analysis across multiple decentralized data sources, without exchanging their data.



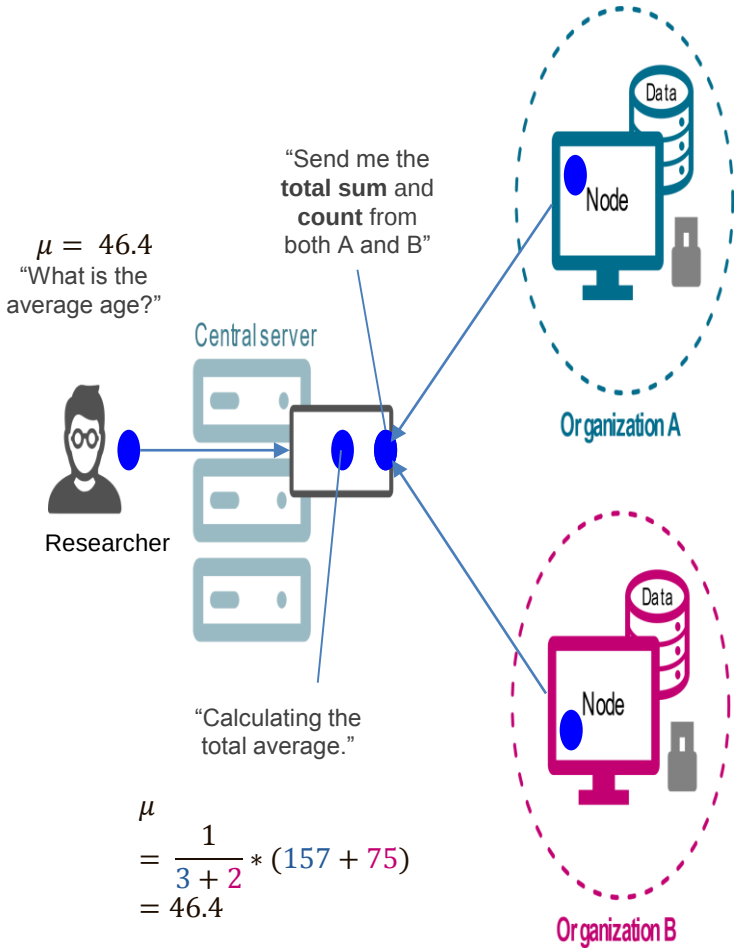
Federated average

$$\mu = \frac{1}{n_a + n_b}\left(\sum_{i=1}^{n_a} \vec{x}_{a,i} + \sum_{i=1}^{n_b} \vec{x}_{b,i}\right)$$

$\vec{x}_a = [57, 68, 32]$

$\sum_a = 157$

$n_a = 3$



$\vec{x}_b = [47, 28]$

$\sum_b = 75$

$n_b = 2$

$$\mu = \frac{1}{3+2} * (157 + 75) = 46.4$$

Synthesis

- Federated approach
- Secure processing environment
- Privacy by design and by defaut
- Data governance layer
- Common data model

Thank you

Principal investigator: Annalisa Trama
annalisa.trama@istitutotumori.mi.it