# Accelerating access to data and research collaborations through trusted research environments (TREs)

**Prof. Jean Louis Raisaro, PhD**

Assistant Professor in Biomedical Informatics and Data Science

Biomedical Data Science Center of CHUV and UNIL

# The Unprecedented Surge in Biomedical Data

The current avalanche of biomedical data is revolutionizing our understanding of health and disease, presenting both **unprecedented opportunities** for medical advancements but also **many challenges**

By 2025, the healthcare industry is projected to generate 2,314 exabytes of data.
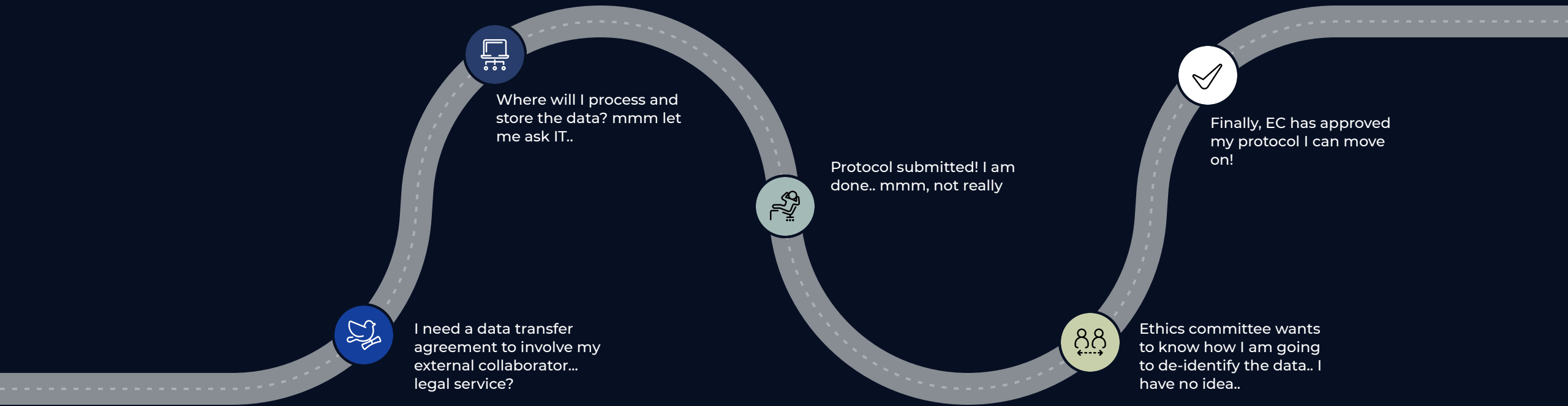(1 exabyte = 732 billions of SD movies)

# This is Marie

- Chief Resident and **clinical researcher** in the Department of Oncology

- She wants to **study overall survival** (OS) of patients affected by brain tumors who undergo radio therapy or pharmacological treatment

- She also wants to **collaborate** with a friend who is professor in AI at the nearby technical school

- They want to develop a **new ML algorithm** that could help clinicians to predict OS for each patient more accurately
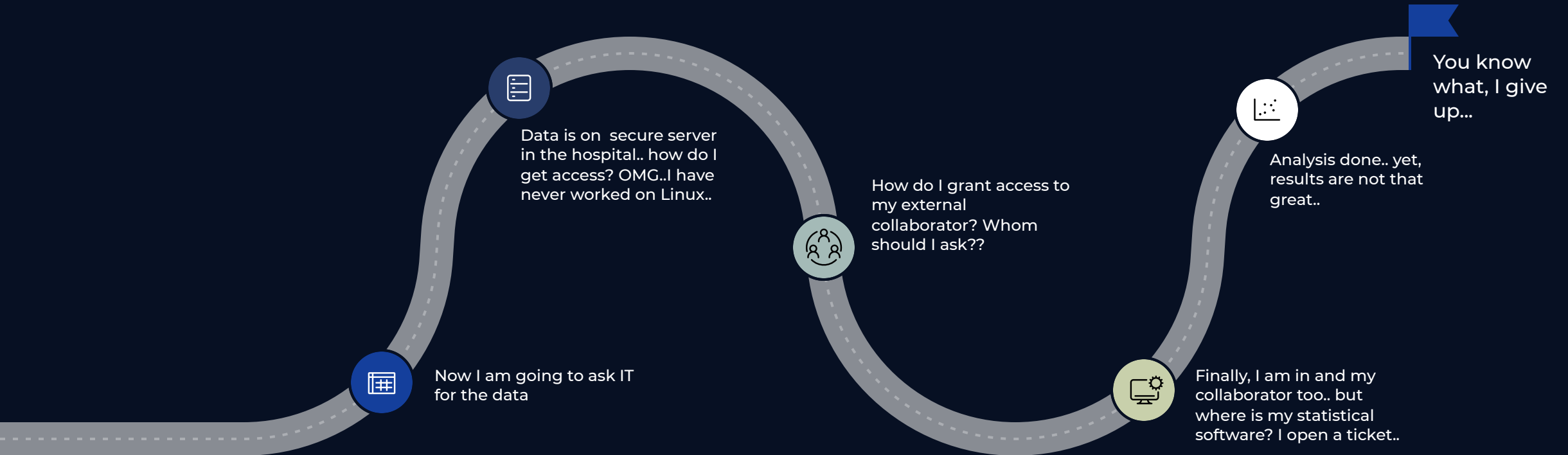
# Marie's painful road to data

Initial idea

Is my study feasible?

How do I get access to the cohort explorer?

10 — mmm, I only get a patient count, that's not very useful..

it seems I will have to wait to know more..

I will write a protocol for ethics so that I can directly get access to the data

Marie's painful road to data (continued)

If you were Marie, would you still be motivated to do research in this setting?

# Today it can take up to several months from an idea to the data

DAYS          HOURS          MINUTES          SECONDS

But researchers are not the only victims…

# This is Paul

- He is a data engineer, working in the hospital research IT and his job is to **help researchers** like Marie to navigate the complex path towards data

- But, Paul's team is **heavily understaffed** and researchers have been waiting for more than 2 months now.. he cannot go faster with such few ressources

- Paul decides to build a new app to enable researchers to do cohort exploration on their own, but **how to let researchers know that the app exists?**

# This is Viktoria

- She works at the DPO office and she is responsible for **data governance**

- She has to make sure that the hospital **complies with all regulations** when it comes to re-use of data for research

- But, she has no way to know what projects are ongoing in the hospital at a given point in time, where is the data for each project and who has access to it.. **she is blinded**..

# Current pain points

### Fragmentation
Research tools and services are fragmented, even within a single institution
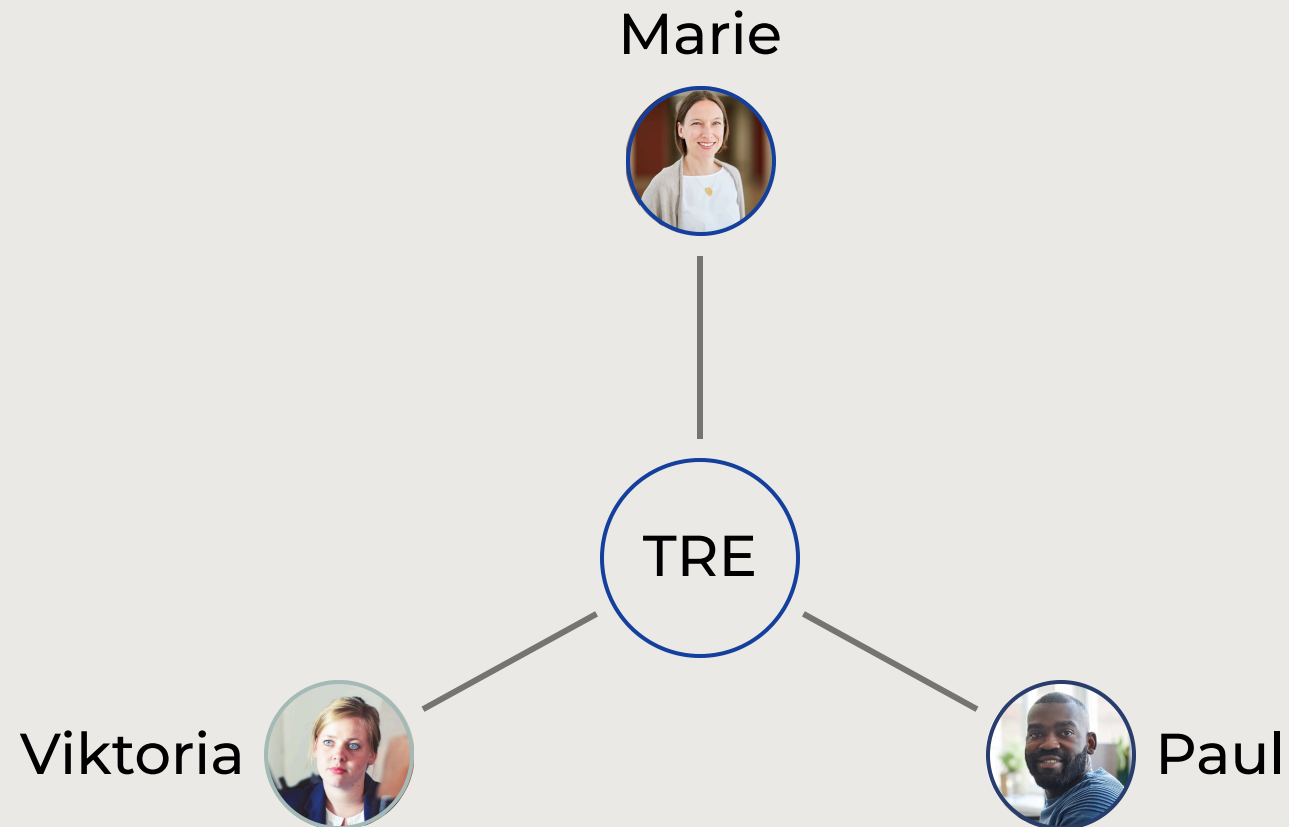
### Lack of collaboration
Researchers struggle to share data and findings with external collaborators

### Compliance difficulties
Meeting regulatory requirements for data privacy and security is hard

# Can Trusted Research Environments (TREs) be part of the solution?

# What is a Trusted Research Environment?

A TRE is a secure computing environment that provides remote access to health data for approved researchers

## TREs make research safer

Personal health data is accessed in a secure and privacy-preserving way

Viktoria

## TREs help make research efficient and collaborative

Researchers can find rich data and generate deep insights to improve healthcare and patients' lives

Marie

## TREs are one-stop shops for data and analytical tools

The data and analytical tools are all in one place

Paul

Marie

# How is data safeguarded?

**Safe People**
Only trained and specifically accredited researchers can access the data

**Safe Data**
Data has been de-identified to protect privacy

**Safe Projects**
Data is only used for ethical, approved research

**Safe Outputs**
All research outputs are checked to ensure no sensitive information leaves the TRE

**Safe Settings**
The data never leaves the TRE

The "Five Safes" model

# Our vision: CHORUS



- CHORUS is a secure virtual platform for **collaborative research**;

- It is cloud-friendly and **infrastructure agnostic**;

- It is going to be **open-source**

- It will be available in **2025**

CHORUS: set of measures of a theme providing the harmonic framework for improvisations

# CHORUS core principles

## Privacy-enhancing technologies

Privacy-enhancing technologies enable a controlled and progressive access to data

## Policy-based sand-boxing

Have clear and transparent policies around data access and use

## User-centric design

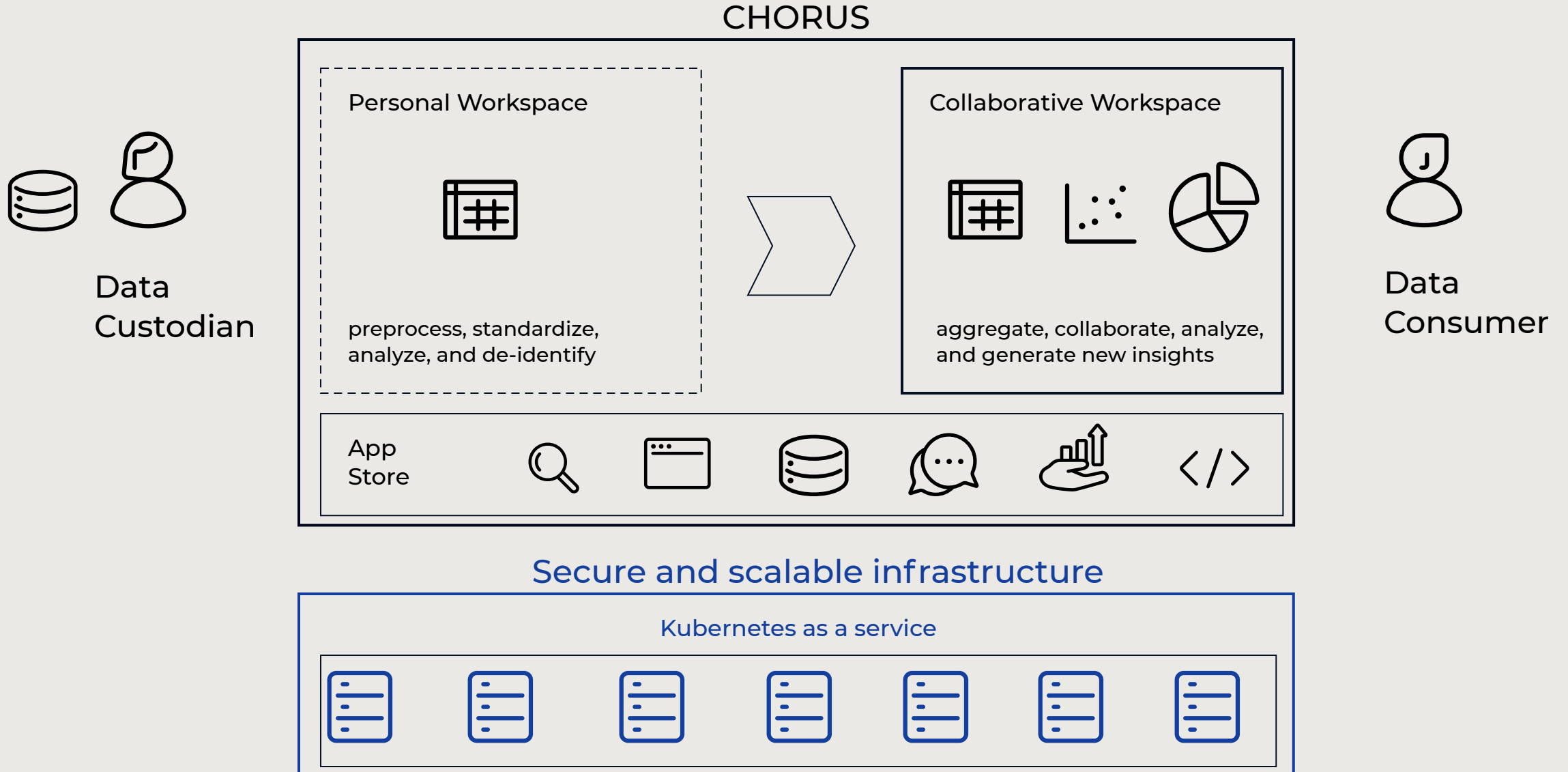Involve users in the design process to build trust and tailored UI/UX
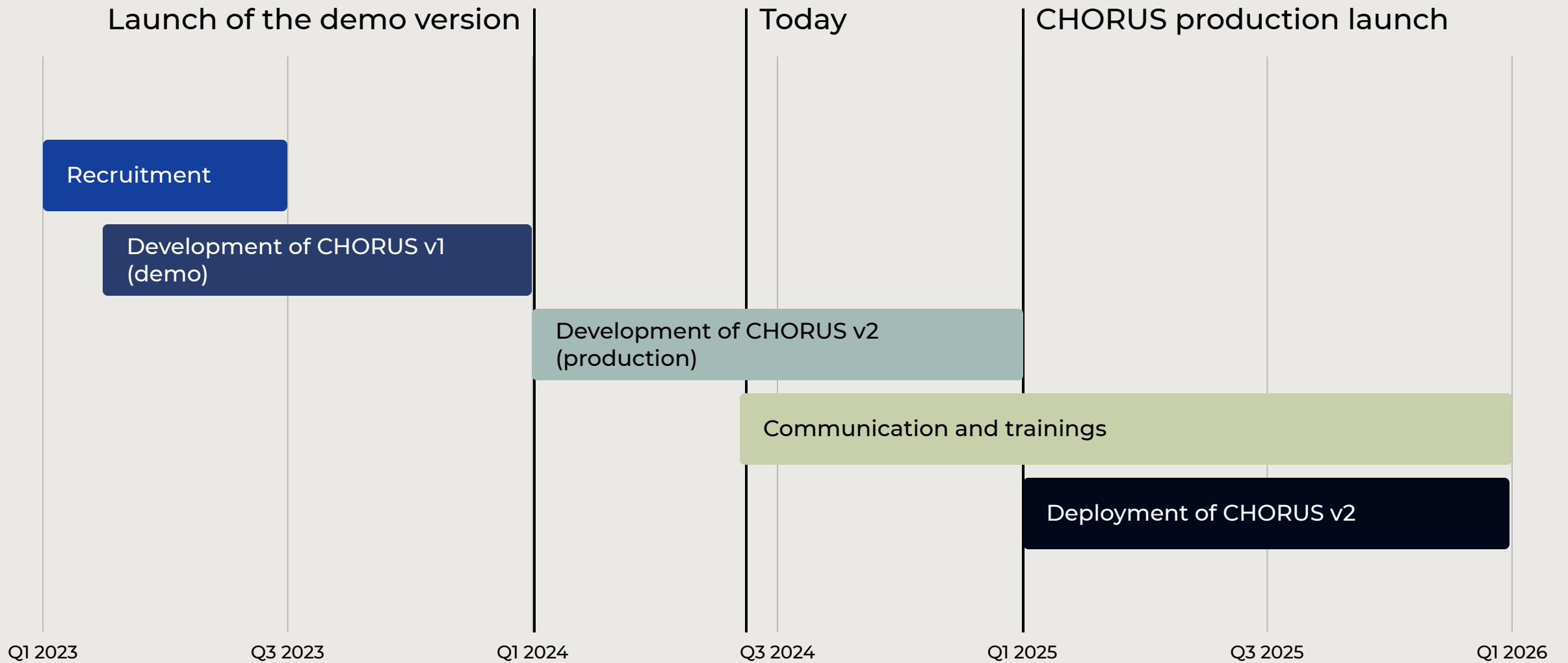
## Open community

Engage with the community to co-design policies, develop new apps and build trust

By focusing on privacy, user needs, community engagement, and transparency, we can build trusted research environments that meet user expectations.

# CHORUS working principle

## CHORUS

**Personal Workspace**

preprocess, standardize, analyze, and de-identify

**Collaborative Workspace**

aggregate, collaborate, analyze, and generate new insights

App Store

Data Custodian

Data Consumer

## Secure and scalable infrastructure

### Kubernetes as a service

# Project plan and timelines



Launch of the demo version

Today

CHORUS production launch

Recruitment

Development of CHORUS v1 (demo)

Development of CHORUS v2 (production)

Communication and trainings

Deployment of CHORUS v2

Q1 2023 · Q3 2023 · Q1 2024 · Q3 2024 · Q1 2025 · Q3 2025 · Q1 2026

# CHORUS at CHUV-UNIL

- Ongoing effort to deploy CHORUS at CHUV as central component of its clinical research platform

- It will allow to safely expose CHUV data to CHUV-UNIL researchers and their external collaborators (including from industry)
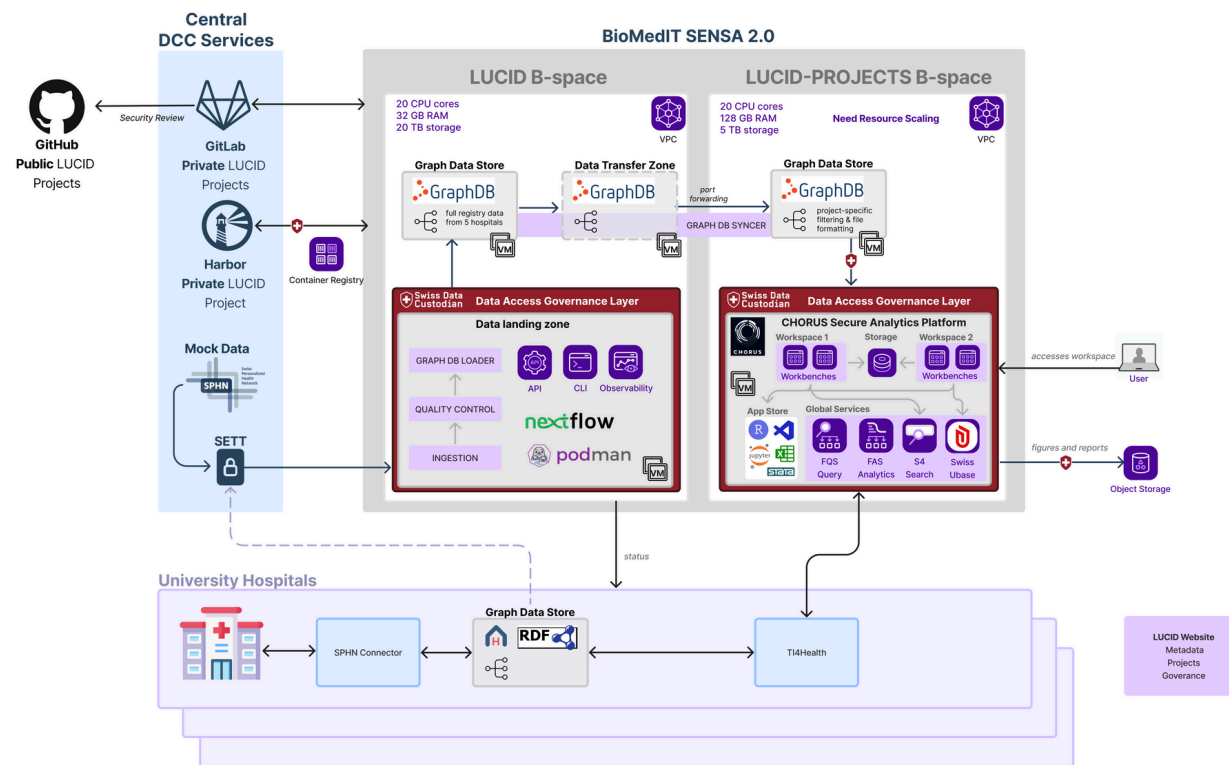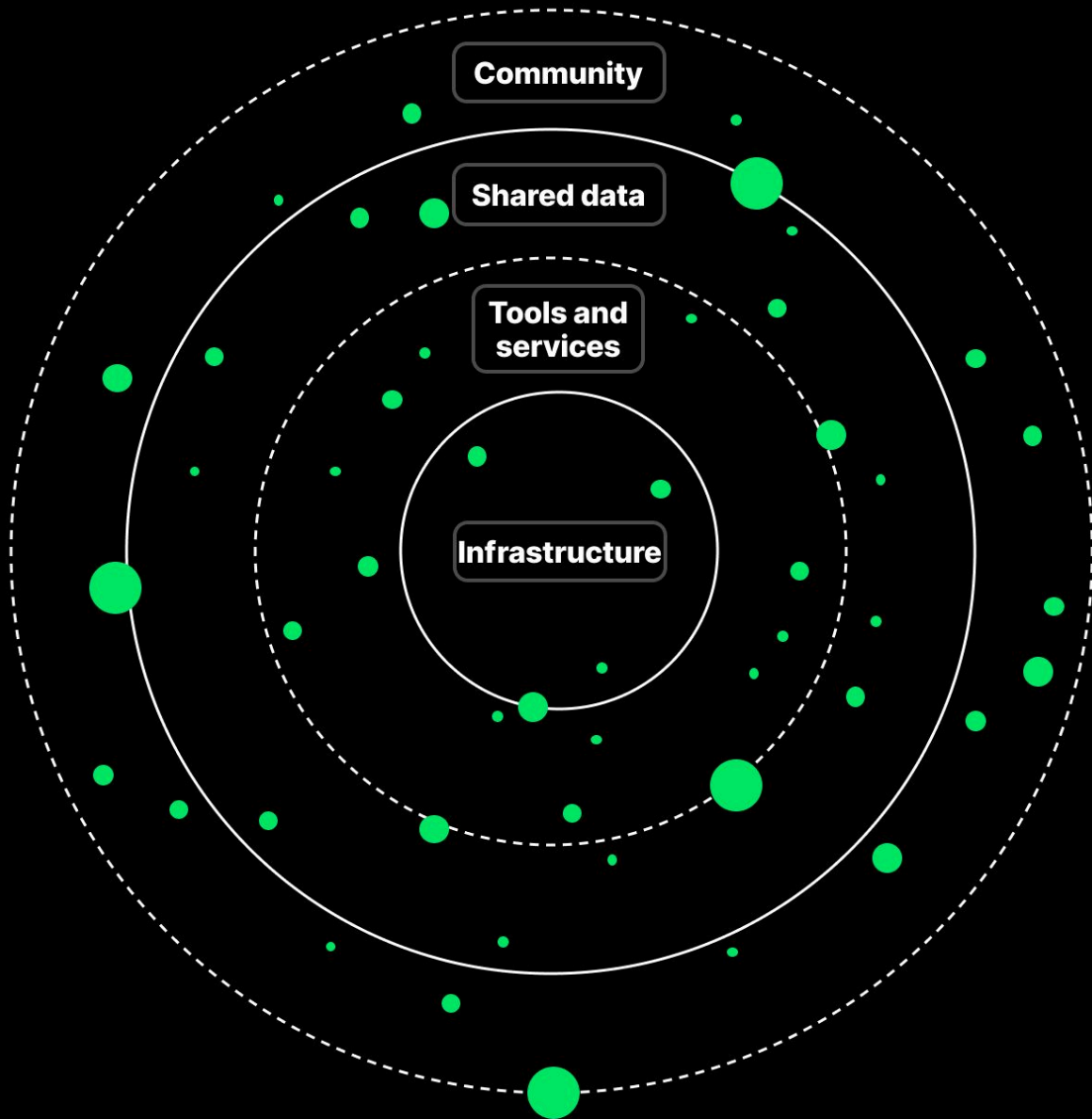
UNIL | Université de Lausanne

CHUV

Faculty of Biology and Medicine

# CHORUS in Switzerland

- Ongoing effort to deploy CHORUS on SENSA 2.0 (powered by Switch)

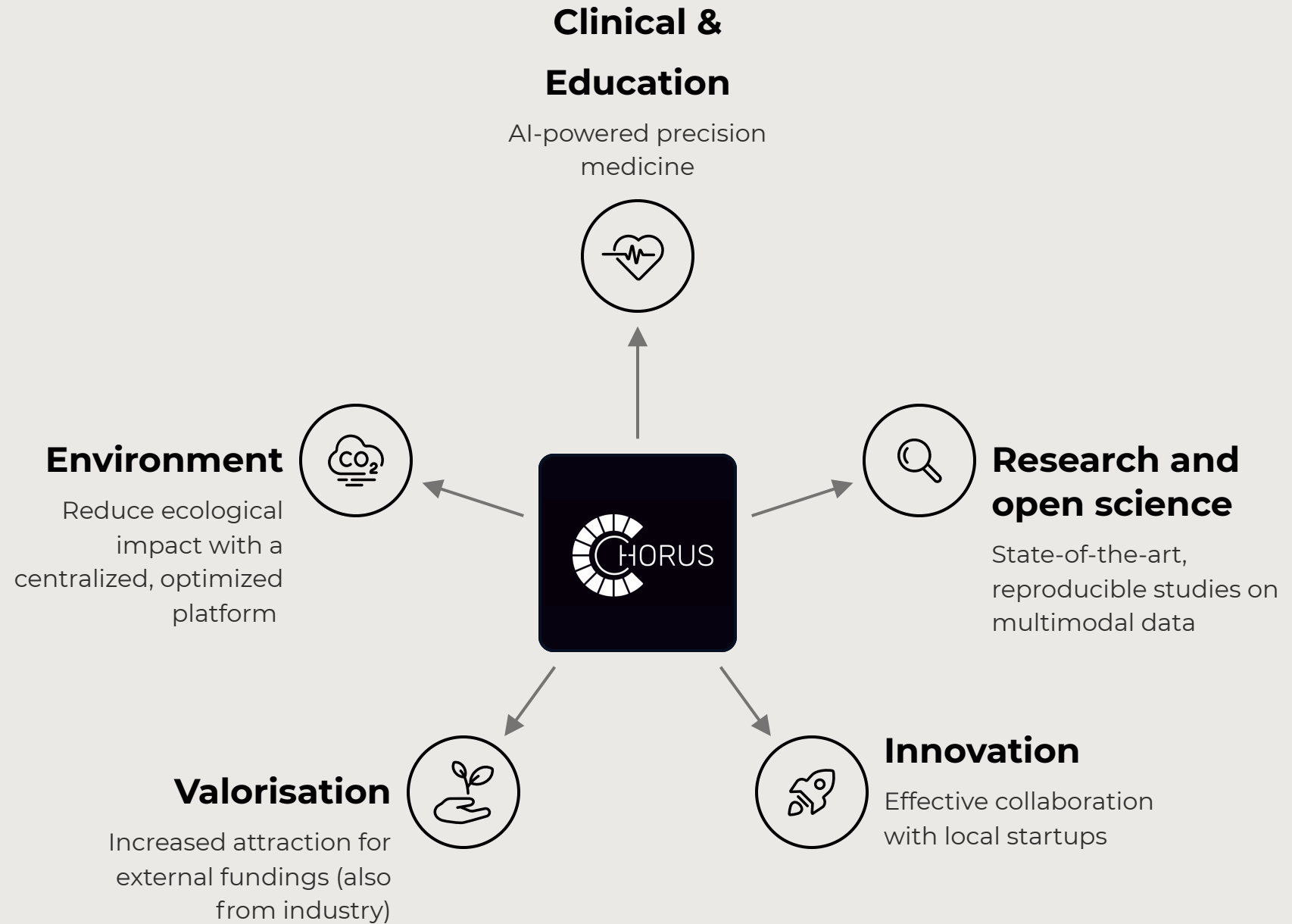- Pilot use case with LUCID National Data Stream

# CHORUS in Europe

- Ongoing effort to install CHORUS in Switzerland, Italy and France as part of the Human Intracerebral EEG Platform - HIP

- 23 institutions already onboarded

The potential impact is enormous

Clinical & Education
AI-powered precision medicine

Environment
Reduce ecological impact with a centralized, optimized platform

Research and open science
State-of-the-art, reproducible studies on multimodal data

Valorisation
Increased attraction for external fundings (also from industry)

Innovation
Effective collaboration with local startups

CHORUS

# Demo

"In an era of increasing digitalization in healthcare, trusted research environments that protect health information are not just an option, they are a necessity."

# Acknowledgments

## Project leadership

Jean Louis Raisaro
Project lead
BDSC

Solange Zoergiebel
Project co-lead
DSI

Nathalie Casati
Tech lead
BDSC

Roberto Colotti
Project Manager
BDSC

## Sponsorship and strategic advisory

Mauro Oddo
Director
DIRC

Raphaël Gottardo
Director
BDSC

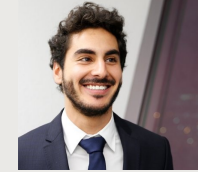Pierre-François Regamey
Director
DSI

Alberto Schena
Director
PACTT

## Biomedical Data Science Center team

Manuel Spuhler
UX/UI specialist

Sami Perrin
Full-stack architect

Bayrem Kaabachi
Data privacy expert

## IT Department team

Pierre-Adrien Ghiringhelli
Data engineer

Yves Jaggi
Data engineer

Samira Najih Besson
Project Manager

## NeuroDigital team

Florian Sipp
DevOps specialist

Julien Dhallenne
DevOps and security specialist

Yoan Blanc
DevOps and back-end specialist

Viktoria Bastic Schmid
Product manager