



Protect-Child

Overview

Eugenio Gaeta



This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement N° 101137423



Pediatric Transplant Challenges and Initiatives

Background and motivation

- **Critical Need:** Transplants meet <10% of global demand; ~20 patients die daily waiting.
- **Pediatric Transplants:** Often the only curative option for end-stage diseases; >5,000 annual procedures in Europe (SOT and HSCT).
- **Challenges:** High treatment-related mortality, long-term complications, immunosuppression issues.

Solutions Needed:

- Find Reliable biomarkers
- Personalized immunosuppression (IS) strategies
- Advanced data integration for complex case management

Future Directions:

- Novel data management approaches
- Enhanced privacy and governance
- Integration of genomic data

Objective

Pilot study to Improve Treatment and Follow-up of transplanted children:

- 200 patients across 4 European hospitals of ERN TransplantChild, plus genome and methylome studies.
- Develop better therapies, reduce treatment toxicities, and improve follow-up care. Focus areas: pharmacogenetics, immune response to immunosuppressants, infections, and epigenetic markers (liver and kidney transplants).

Build Cross-Country Data Infrastructure:

- Enable secondary use of clinical and genomic data across Europe for pediatric transplants.

Technologies and Infrastructure:

- Leverages ELIXIR, GDI, Quantum Computing, and AI for federated data analysis.
- Follows EHDS, EHDS2pilot, and TEHDAS standards for quality, privacy, and security.

Ambition: Create a European reference model for integrating clinical and genomic data to transform pediatric transplant care and research.

Protect-Child



Concept and approach

Goal: Develop a federated data space for the TransplantChildERN enabling secondary use of health and genomic data across Europe.

Approach: Build a **multi-source, multi-standard** infrastructure aligned with **EHDS** principles, leveraging **GDI** and **ELIXIR** synergies.

Core Pillars:

- **Infrastructure:**
 - Containerized, cloud-based architecture.
 - Secure Processing Environments at each center enable federated access.
 - EHDS-compliant services for data discovery and access.
- **Data Governance:**
 - Shared governance model across all centers.
 - Blockchain-supported digital workflows for access requests, ethics approvals, and consent.
- **User-Friendly Assistants:**
 - Tools for researchers, clinicians, innovators, and policymakers.
 - Designed to simplify access to cross-border, multi-source datasets.
- **Design Thinking & Co-Creation:**
 - Involves stakeholders in the design process.
 - Starts with pilot use cases in renal and liver transplants, scaling to full ERN coverage.



The project focuses on studying two types of rare diseases in children that require organ transplants: **liver and kidney disorders**. These diseases are linked to specific genetic issues, which create unique challenges for diagnosis and treatment. The project will examine **200 patients** who have had liver or kidney transplants, **analyzing their genetic material and certain chemical markers in their DNA to find new genes and regions in the genome associated with these diseases**.

The methodology will involve

- Whole Genome Analysis,
- Polygenic Risk Score calculation, and
- Methylome studies using episignature analysis.

The project will also incorporate the genomic data into the Genomic Data Infrastructure (GDI).

Key Goals:

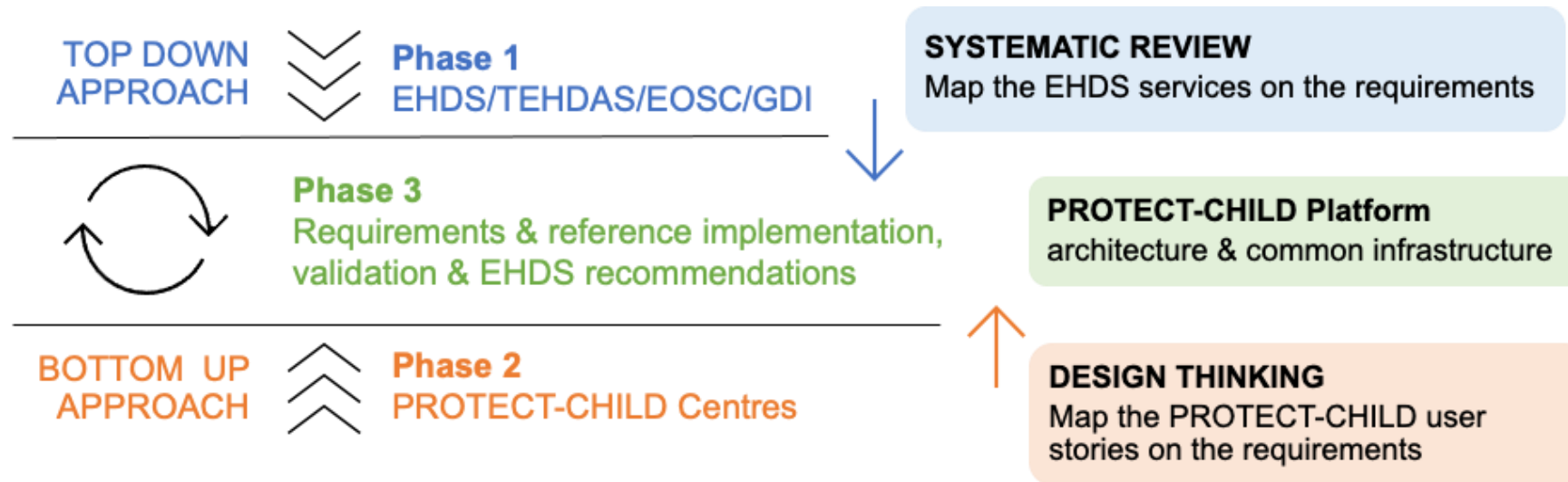
- **Enhance Genetic Studies (GWAS):** The pilot aims to improve large-scale genetic studies by adding data from around 1 million specific genetic variations (SNPs) obtained from comprehensive genome sequencing.
- **Develop Better Risk Prediction (PRS):** Using advanced machine learning and genomic technology, the pilot will create a better tool (Polygenic Risk Score) to predict the risk of disease for liver and kidney transplant patients based on their genetic makeup.
- **Find New Genetic Markers:** The pilot will identify new genetic markers that can predict how patients respond to medications, including potential adverse reactions and drug toxicity.
- **Understand Genetic Susceptibility:** Researchers will look for genetic markers that show how susceptible patients are to drug reactions, viral illnesses, and environmental factors.
- **Study Epigenetic Markers (Methylomes):** By examining specific chemical changes in DNA (methylation), the project will identify groups of patients with similar DNA markers and discover new disease-related signatures.
- **Incorporation into a Larger Initiative:** The genetic data collected will also be included in the Beyond 1 Million Genomes (B1MG) project, a larger initiative aimed at understanding genetic diseases.

Protect-Child

Co-creation approach

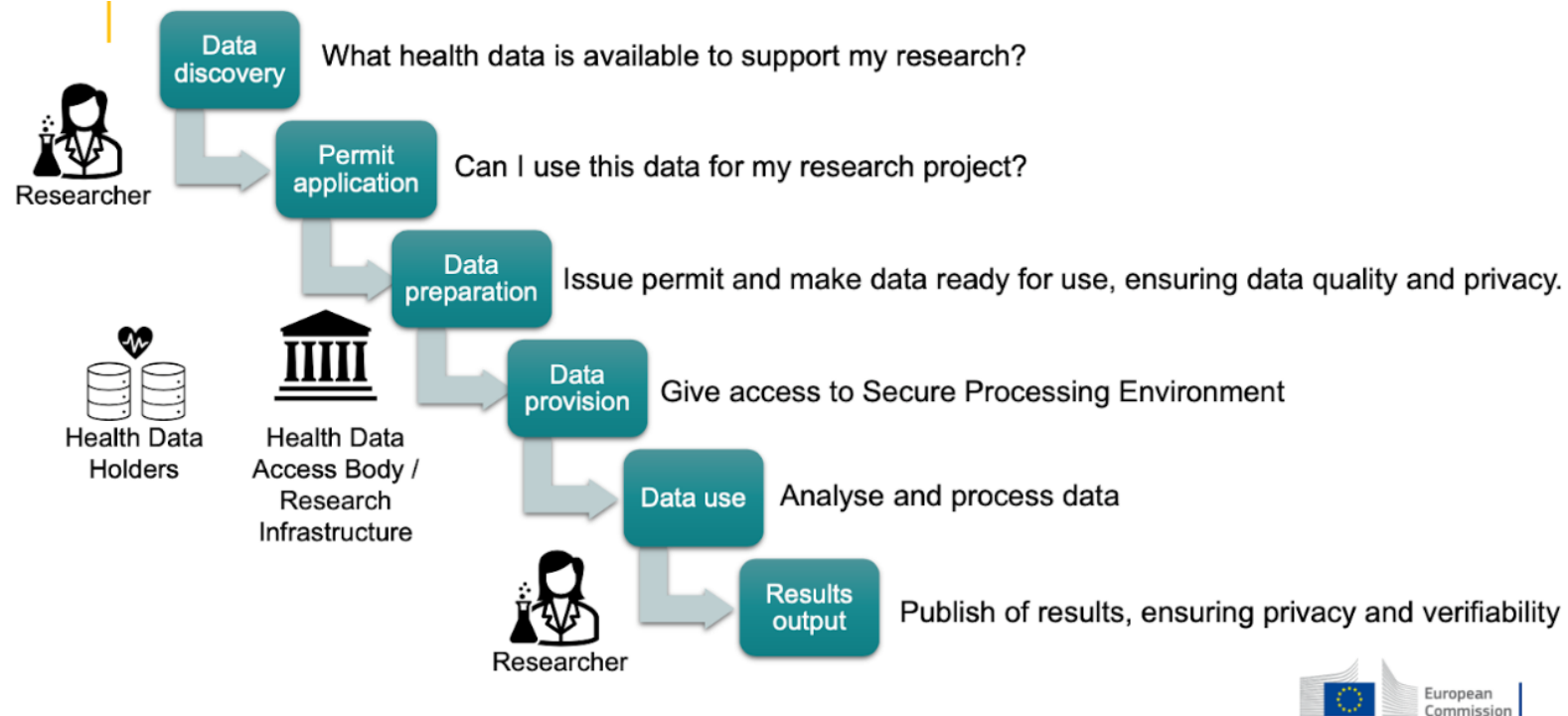


Space methodology



In PROTECT-CHILD, the SPACE methodology combines **top-down analysis of European initiatives** like EHDS and GDI with **bottom-up stakeholder engagement to refine requirements**. **Socio and cognitive science methods are employed** to ensure feasibility, interoperability, and compliance with evolving regulations, resulting in a comprehensive approach to defining the project's data space.





The PROTECT-CHILD concept as IDEA4RC provides an architecture for **implementing the European Health Data Space (EHDS) user journey**, which includes the six phases of secondary use of data: **data discovery, permit application, data preparation, data provisioning, data use, and results output**. PROTECT-CHILD enables **secure and ethical access to health data for research and innovation**. The pilot includes also primary use of data for the collection and analysis of genomics data. PROTECT-CHILD is a borderline use case of EHDS because it merges primary and secondary use of data.

Technological solution



EHDS Capsule: Secure and Interoperable Healthcare Data Ecosystem

A specialized environment for local privacy-preserving data processing.

Key Features:

- **Structure unstructured data:** Use biomedical NLU to extract and transform narrative information into structured EHRs.
- **Enhance data reuse multi-standard data collection:** Unify data capture in OMOP and FHIR on a common data model ensuring interoperability at different level for primary and secondary use of data (OLTP system).
- **Local analysis :** analytics software stack, enable local data processing, and support massive data processing (OLAP system).
- **Federated learning and multi-party computation:** Enable collaborative data science without central data aggregation, ensuring privacy and security and first effort for integration of quantum computing.
- **Scalable deployment:** Decouple infrastructure, enabling on-premises, public, private, or hybrid cloud environments.

Benefits:

- Enable real-time data analysis and AI model training preserving privacy
- Support interoperable data models and services using multi-standard data transformation hubs (HL7 FHIR, OMOP, etc.)
- Address EU Data Strategy's topic on GDPR- and EHDS-compliant Secure Processing Environment service



Technological solution

EHDS Capsule architecture

Zero Trust environment implemented with service meshes

CLUSTER

CAPSULE MESH

OLTP services

OLAP services

Governance services

Data discovery services

Beacons and genomic services

CONTROL PLANE

INGRESS/
EGRESS
GATEWAY

mTLS

+ Privacy
+ Isolation

+ Interoperability

+ Security
+ Trust

+ Data Quality

mTLS

DATA PREPARATION

ETL engine

NLP engine

Quality checks

QUANTUM COMPUTING

Novel Quantum algorithms
for genomics and federated
data processing

**Cluster
services**

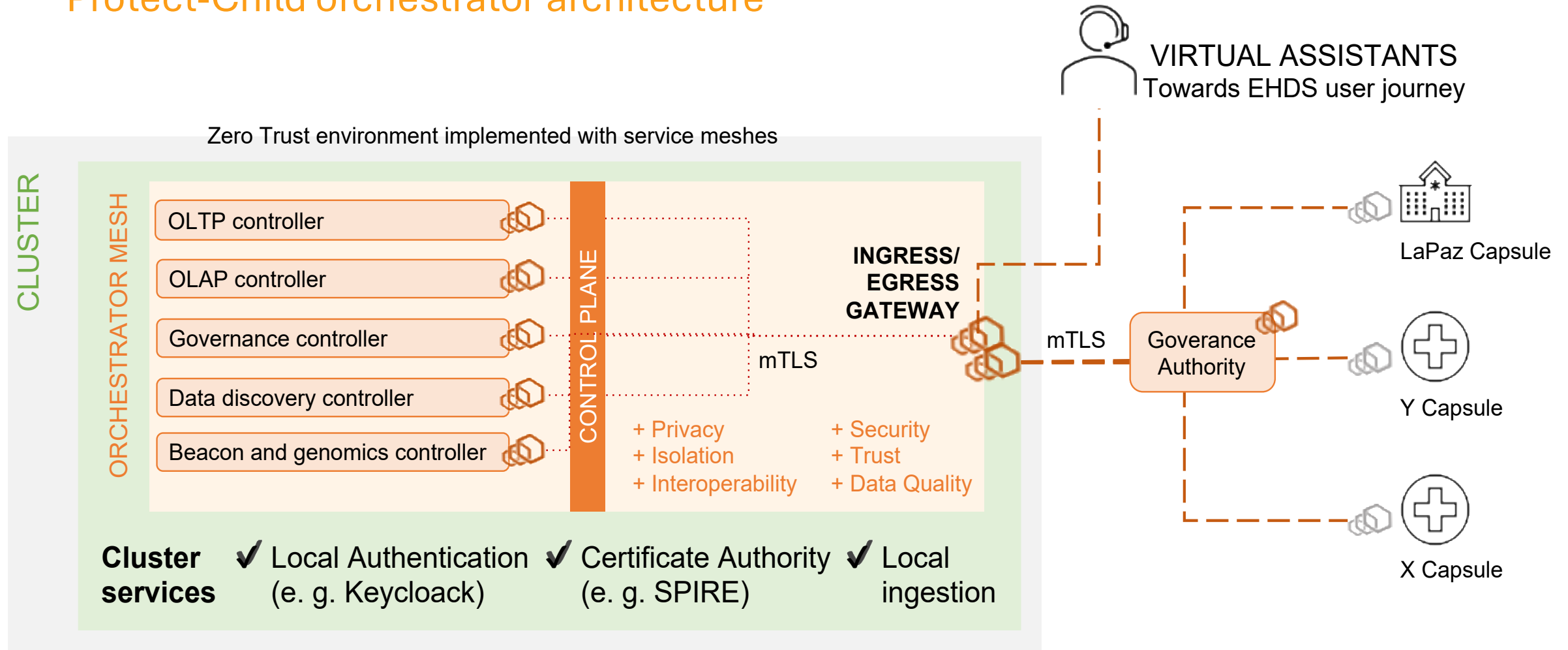
✓ Local Authentication
(e. g. Keycloak)

✓ Certificate Authority
(e. g. SPIRE)

✓ Local
ingestion

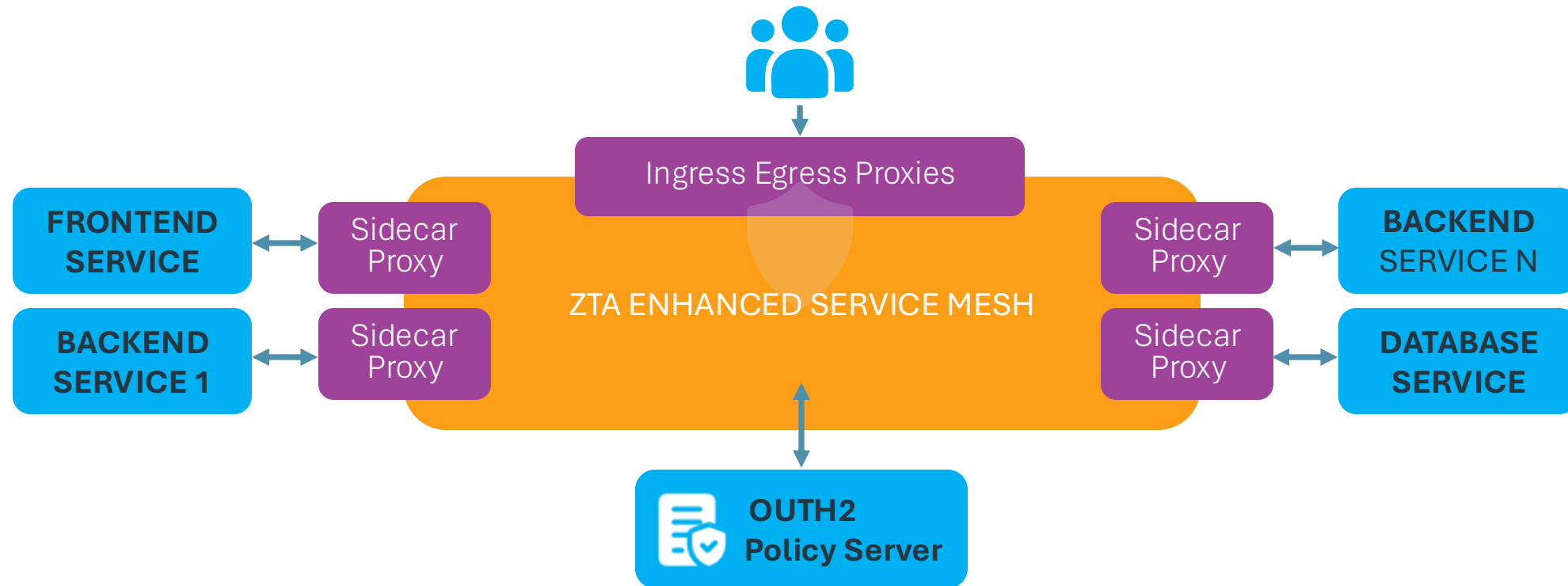
Technological solution

Protect-Child orchestrator architecture



Technological solution

Protect-Child Zero Trust security logical model



Zero Trust turns “implicit trust zones” into **explicit, continuously validated trust decisions**, throttling both external intruders and insider threats while keeping legitimate business flows humming.

Zero Trust: How it Drives “High-Security” Outcomes

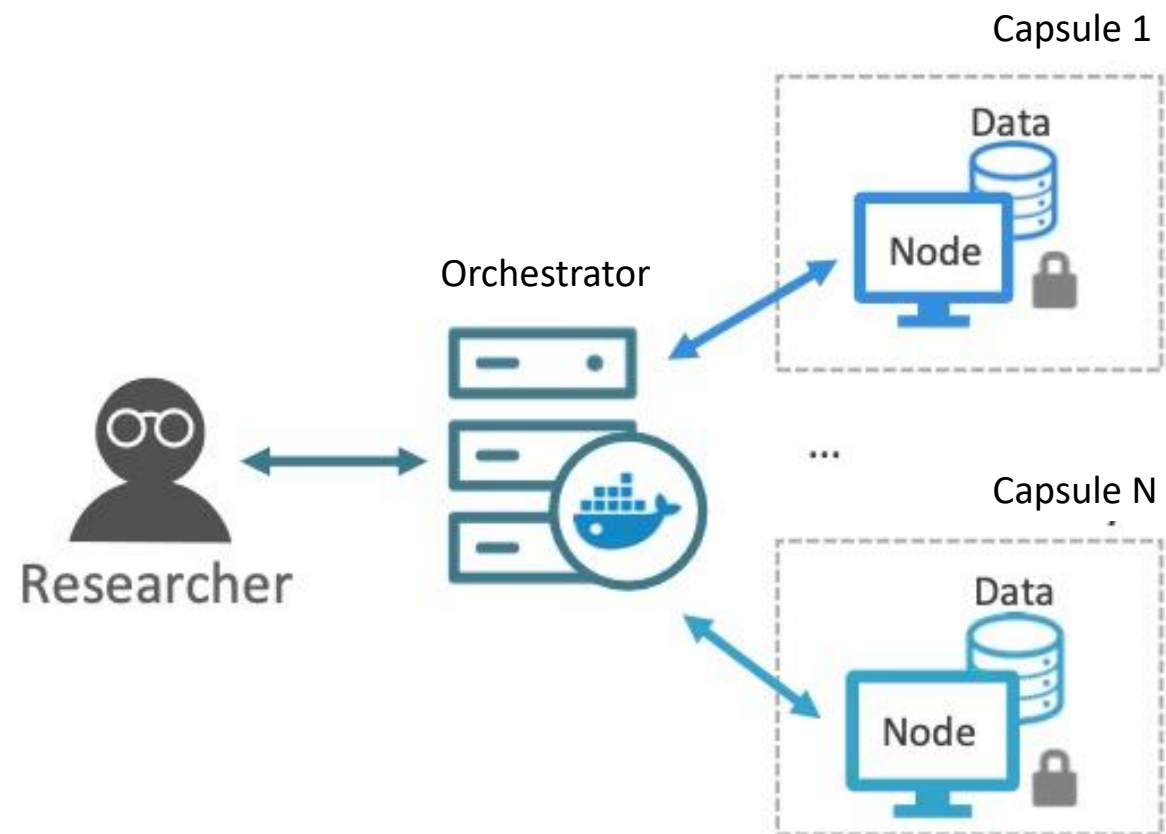


Principle	What It Means	How It Raises Your Security Bar
Never trust—always verify	Every request, no matter the network location, is authenticated and authorized in real time.	Eliminates implicit “soft center” that attackers rely on after the first foothold.
Strong identity at the core	Users, workloads, devices, APIs all have cryptographically validated identities.	Stops credential-stuffing and spoofing; ties every action to a provable entity.
Least-privilege, just-in-time access	Permissions are scoped to the minimum and expire quickly.	Shrinks attack surface and limits the blast radius of compromise.
Micro-segmentation & continuous inspection	Traffic between micro-services is encrypted and policy-checked hop-by-hop.	Blocks lateral movement, insider abuse, and stealthy persistence.
Context-aware, adaptive policy	Signals like device health, geolocation, and behavior score adjust access on the fly.	Detects & contains anomalous actions before data is exfiltrated.
Unified telemetry & automation	Logs, network flow, and threat intel feed ML-driven enforcement loops.	Speeds mean-time-to-detect/ respond (MTTD/MTTR) and cuts manual errors.



Technological solution

Protect-Child federated learning approach



“process-in-place, share-only-what-is-needed.”

Protect-Child federated learning approach



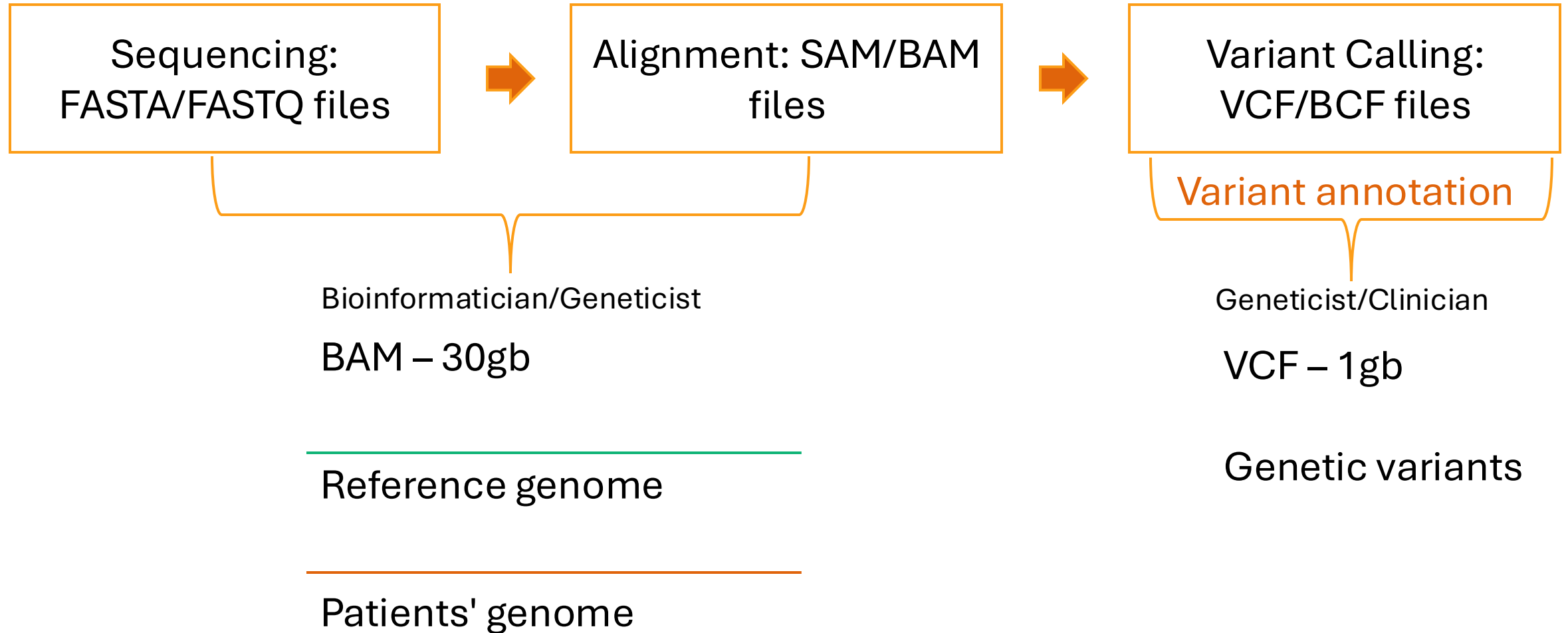
How It Delivers High-Level Privacy

Building Block	What Happens	Privacy Gain
Data never leaves the device / site	Each phone, browser, or hospital runs the training or analytics job locally.	Raw personal data never sits on a central server , eliminating the biggest breach target.
Model-update sharing, not data sharing	Only gradients or parameter deltas are sent to the coordinator.	Updates are far less revealing than the underlying records, shrinking leakage risk.
Secure aggregation	A cryptographic combiner sums the updates so the server sees only an encrypted total , never any one participant's contribution.	Even a curious or compromised server can't read individual updates.
Differential privacy noise	Tiny, mathematically calibrated noise is added before or after aggregation.	Guarantees that attackers can't infer whether any single user was in the training set.
Short-lived keys & TLS tunnels	Every round re-keys; traffic is end-to-end encrypted.	Blocks passive network sniffing and prevents replay attacks.
Edge control & consent	Participants can pause, throttle, or delete their local data at any time.	Aligns with GDPR/CCPA "right to be forgotten" and minimizes regulatory exposure.



Technological solution

Genomic analysis workflow and Massive data management





Thank you!

This project has received funding from the European Union's Horizon 2020
research and innovation programme under grant agreement N° XXXXXX

