

The Alan Turing Institute

Sensitive AI workloads on
national HPC infrastructure: the
FRIDGE satellite TRE model

Matt Craddock

HPC-AI Advisory Council, Leicester, 14/10/26

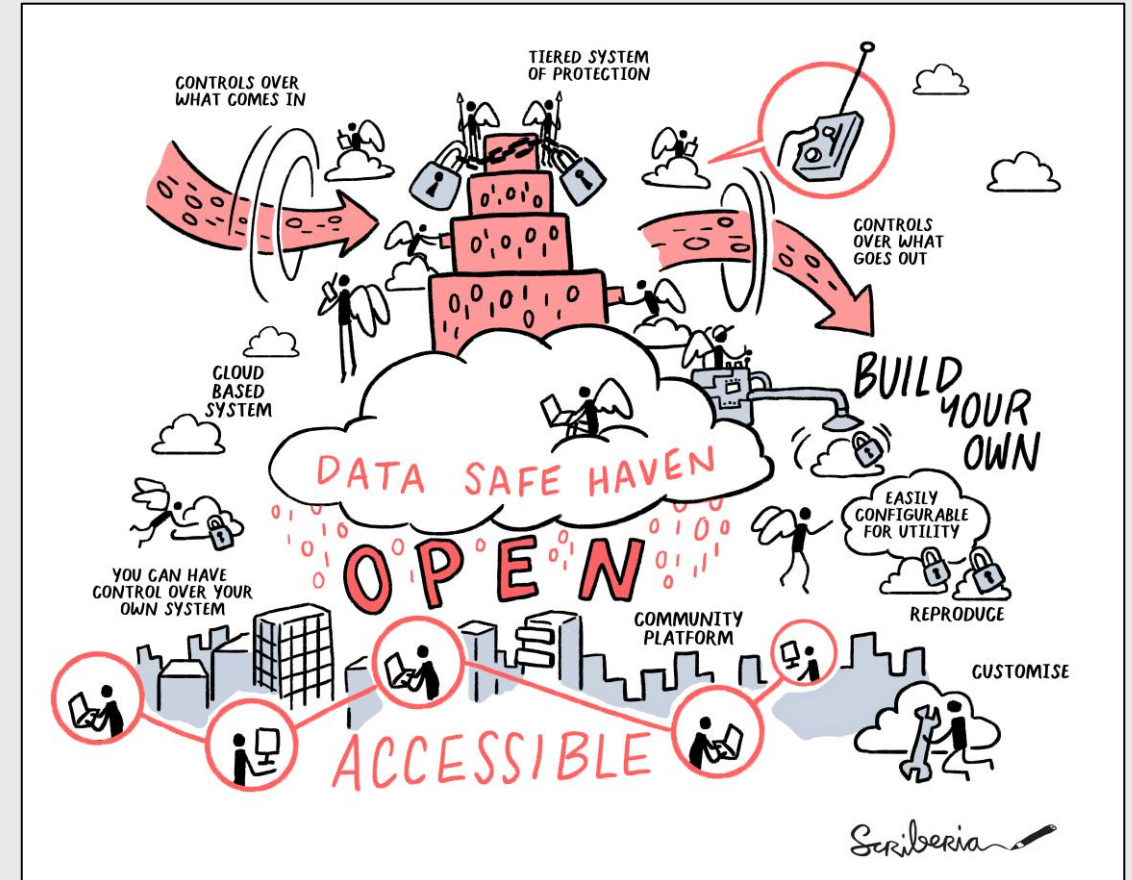


The challenges of research with sensitive data

- Sensitive data can be extraordinarily valuable for research
- Researchers need access to it at scale and in depth
- The potential for misuse is high
 - Identity exposure
 - Commercial exploitation e.g. discrimination by employers or insurers

Trusted Research Environments

- Tightly controlled environments
- Control over who is allowed access
- Control over what they can bring in
- Control over what they can take out



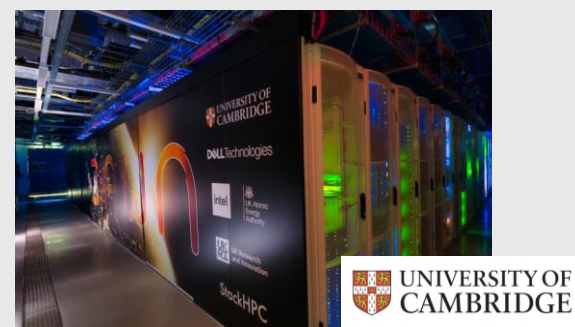
Real projects, real data, real TREs

- **UK Biobank**: Large 500,000 subject longitudinal research cohort. Long follow-up includes many different health outcomes.
- **PharosAI**: Large-scale multimodal data. 12,000 patients per year. Digital pathology, imaging, DNA/RNA, longitudinal clinical data.
- **AI Centre for Value Based Healthcare**: 170,000 brain MRI images from 40,000 patients with metadata and clinical radiology reports.



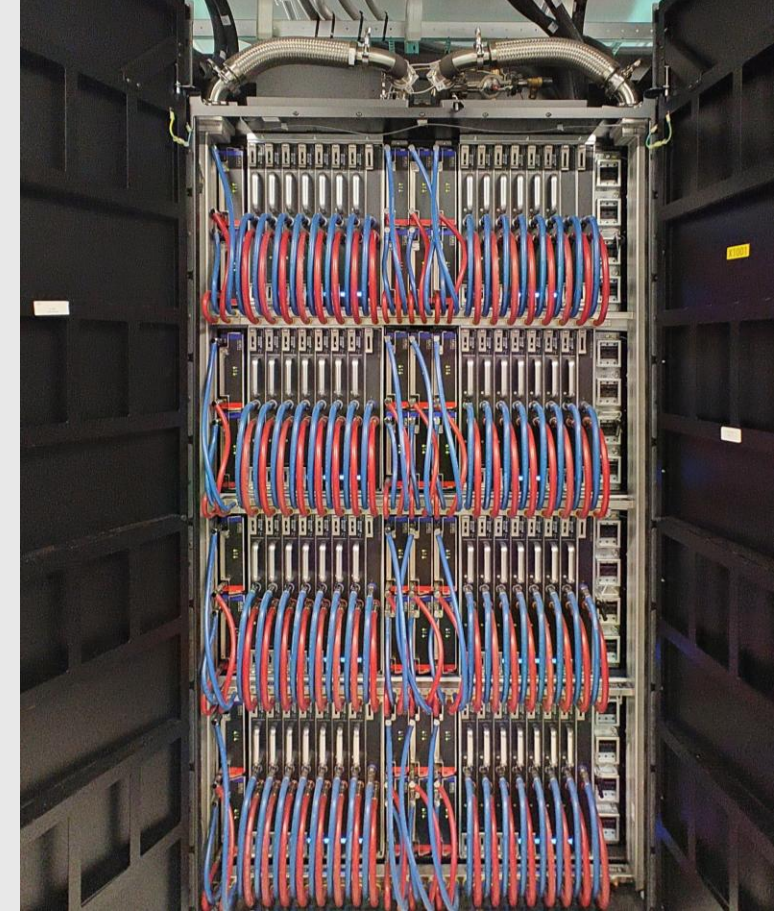
AI Research Resource

- High-performance computing facilities designed for cutting-edge AI research
- Two sites: Isambard-AI (Bristol) and Dawn/Zenith (Cambridge)
- Huge amounts of computing power



Trusted Research Environments on Supercomputers

- Research needs powerful hardware
- AI in particular needs GPUs
- Fantastic national supercomputing resources
- Supercomputers are not TREs
- How do we use supercomputers for trusted research?



FRIDGE



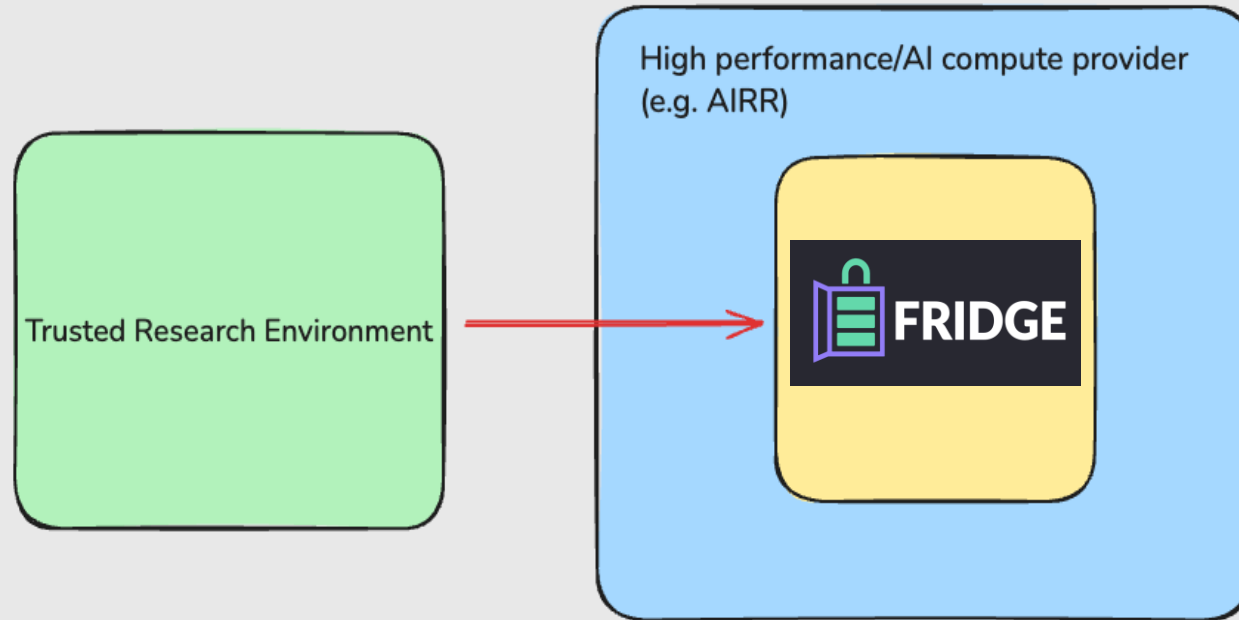
FRIDGE
A DARE UK PROJECT

The Challenge

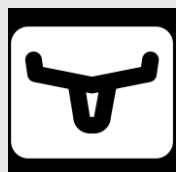
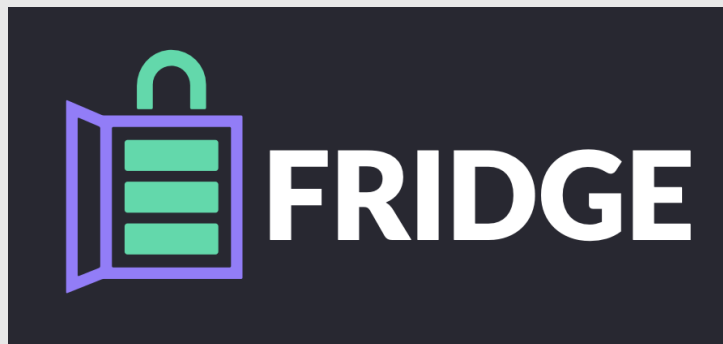
- Creating a **secure enclave** on a **multi-user system**
- **Balancing security burden** between TRE and supercomputers
- Avoiding bespoke **data-sharing agreements** for each project
- **Efficient** use of a shared resource

What is a FRIDGE?

- A **secure enclave** hosted within broader computing infrastructure (e.g. AIRR, Azure)
- An **extension to an existing TRE**, not a standalone TRE
- Both a technical implementation and an operating model based on **shared responsibility**



What's inside the FRIDGE?

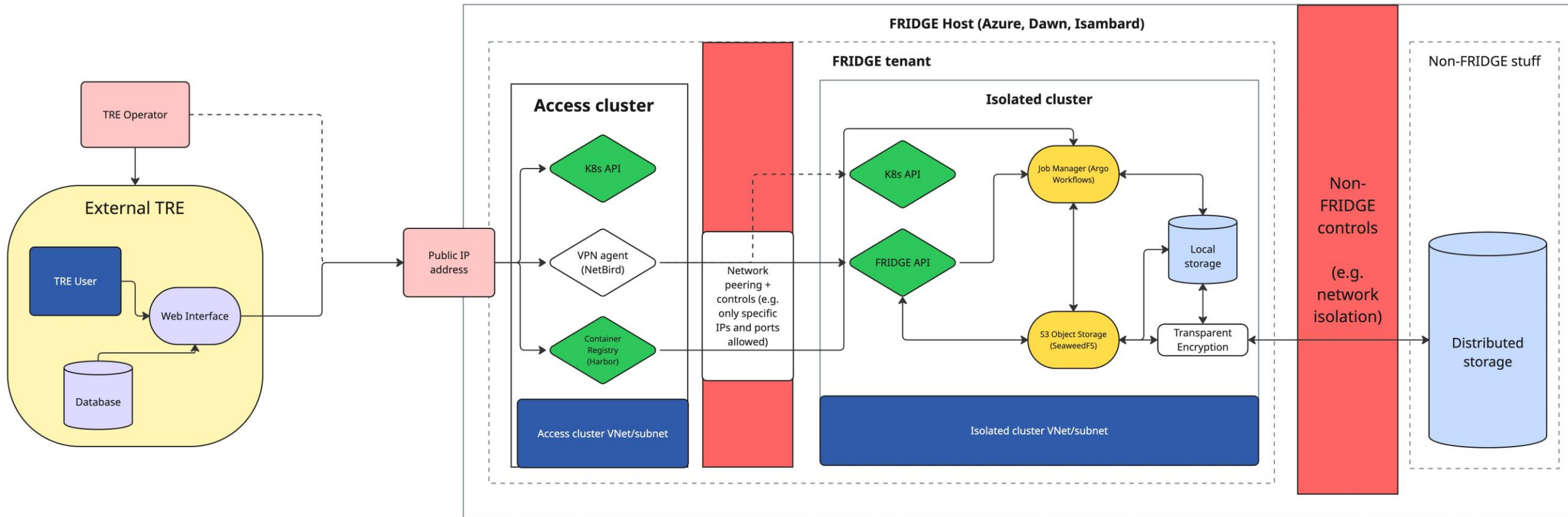


cilium

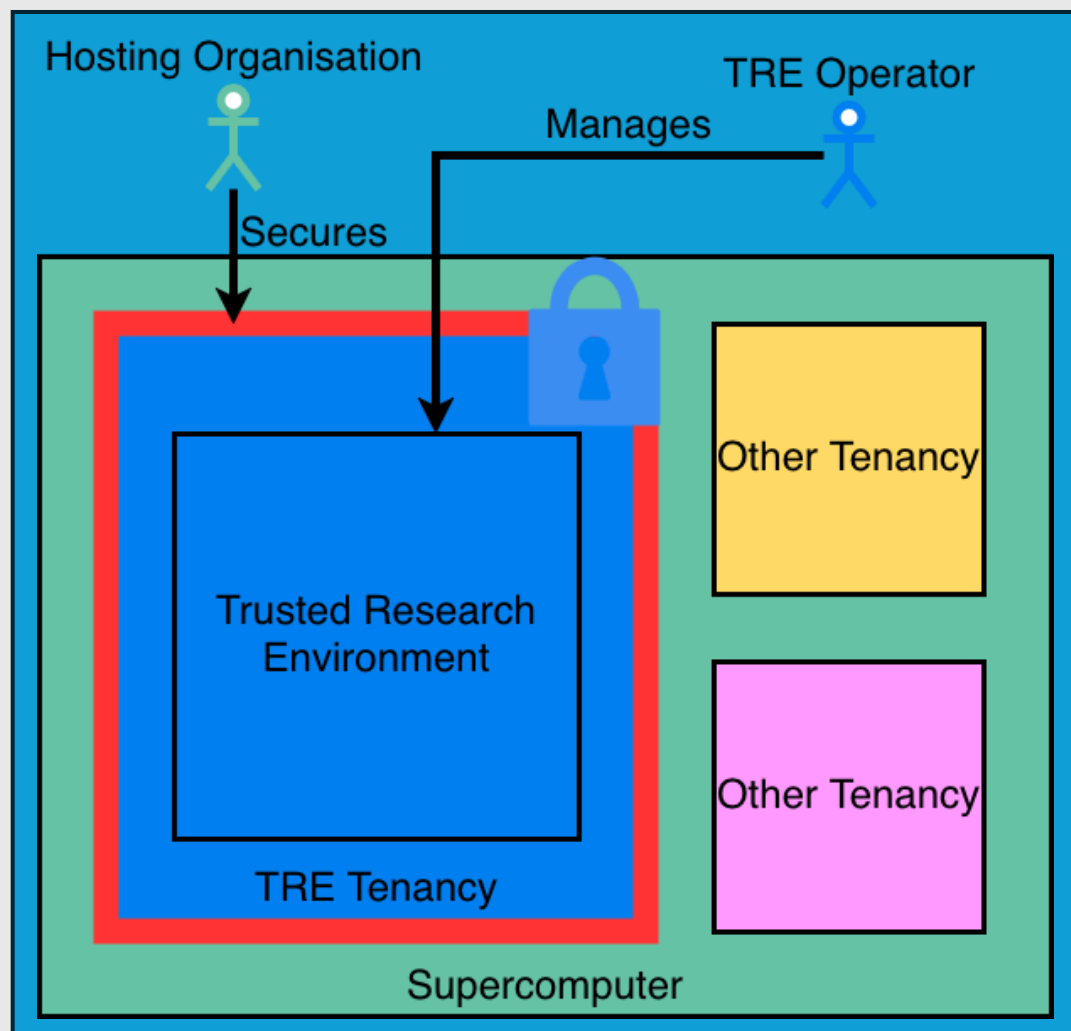
- Kubernetes-based (platform-agnostic...ish)
- VPN-based access
- Workload orchestration (Argo Workflows)
- Object and block storage (SeaweedFS/Longhorn)
- Custom API (FastAPI)
- Container registry (Harbor)
- K8s-native networking controls (Cilium)



The architecture of a FRIDGE



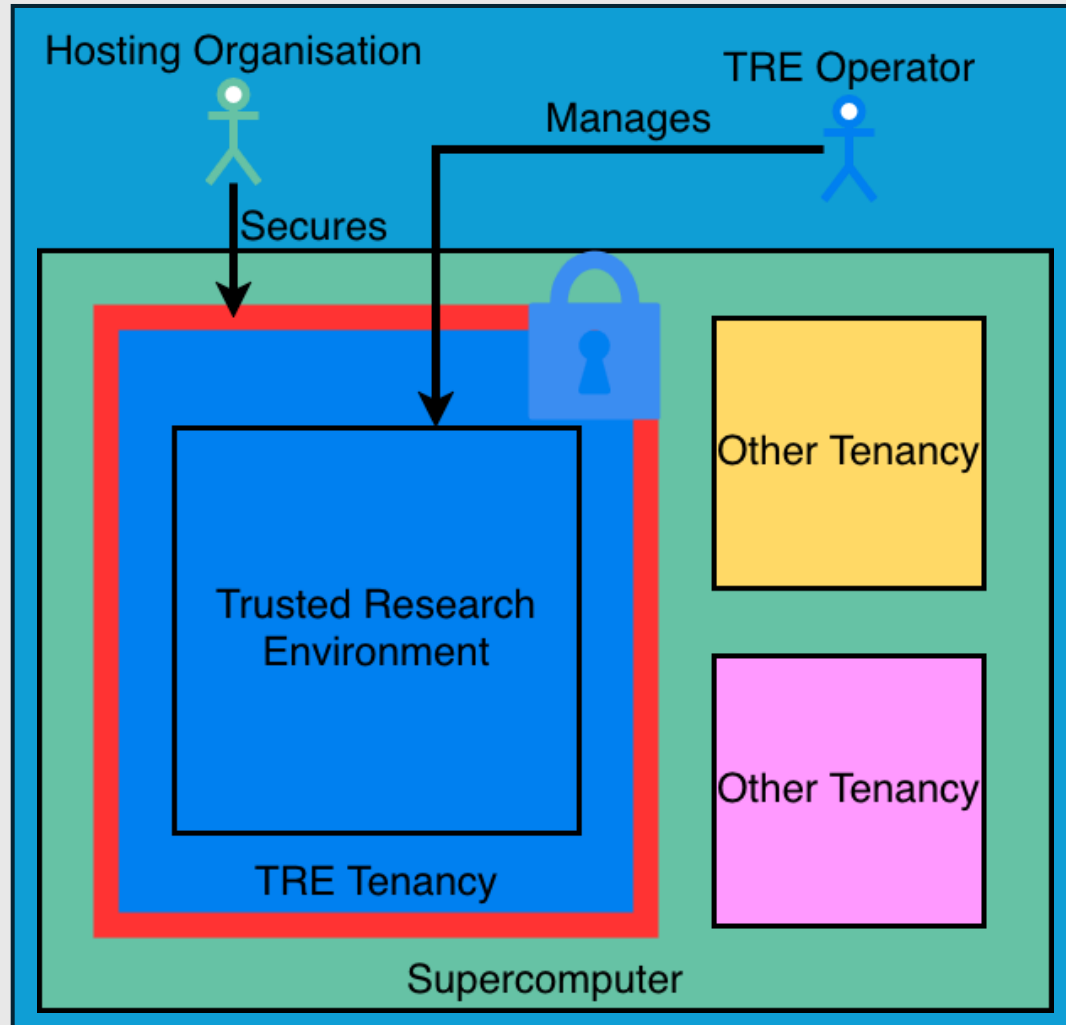
TRE Tenancy & Shared Responsibility



Shared Responsibility

- **Hosting Organisation** secures the boundary
- **TRE Operator** manages **FRIDGE**
- TRE Operator holds **sole responsibility** for governance, access, data processing

TRE Tenancy & Shared Responsibility



TRE Tenancy

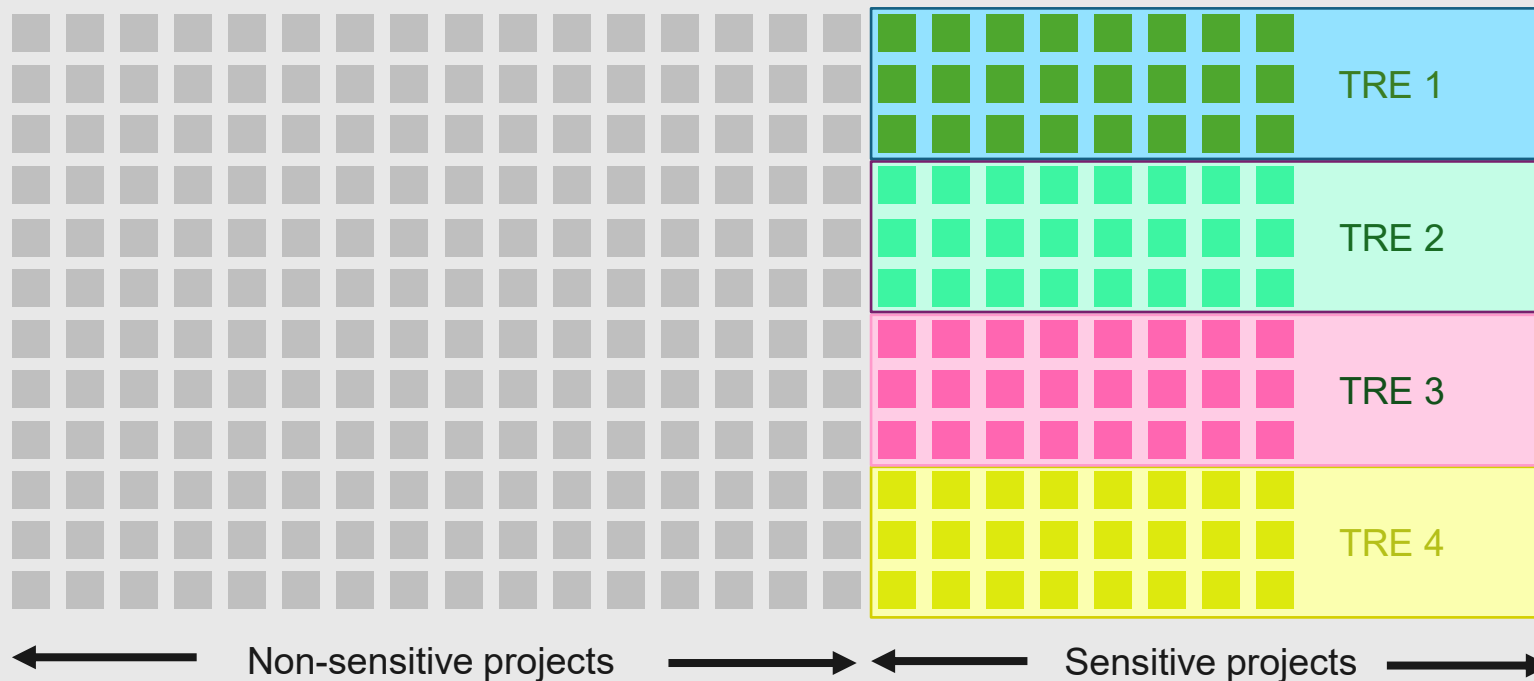
- Secure area, housing K8s
- Network isolation and data encryption
- Set of requirements to be implemented
- Enforced at infrastructure level



- Bridging Research Infrastructure for Data Governance Extension on AIRR!
- Implementing a shared responsibility model
 - UK Biobank, Pharos AI, and AI Centre for Value Based Healthcare have the data
 - KCL/Turing have the TREs
 - AIRR has the hardware
- Real projects, real data, real TREs



AIRR-BRIDGE: Shared responsibility



FRIDGE developed a **satellite TRE** model to **isolate different parts** of a supercomputer for **different TREs**.

AIRR-BRIDGE will **validate** this approach in a **production** environment.

Future directions

- LUKS based encryption (BYOK)
- Off-node storage
- Interfacing with host TRE systems (e.g. KCL SLICE)
- Adapting for inference as well as training

FRIDGE

- Demonstrating how the power of supercomputers can be harnessed by TREs
- Providing a model for how sensitive data research can be conducted on HPC systems
- Open source, with both technical and governance frameworks

<https://alan-turing-institute.github.io/fridge/>



FRIDGE

A DARE UK PROJECT



AIRR-BRIDGE

A DARE UK PROJECT