

# Seeing the FOREST through the TREES

Susheel Varma, Chief Technology Officer, HDR UK

## Trust me, I'm a Trusted Research Environment Evangelist 🗣️

- Recovering pure mathematician with a 17 year meandering career in:
  - Computational Physiology, Health Informatics, Bioinformatics
  - Cardiac Electrophysiology, Orthopedics, Virology, Medical Devices, Dementia, Colorectal Cancer, Rare Disease
  - Technologist, Global Standards and expertise in setting up and operating national and international Research Infrastructures (4)
- Education/Awards:
  - PhD in Computational Systems Biology – Cardiac electrophysiology
  - Exec. MBA in Management of Research Infrastructures
  - Fellow of the British Computer Society
- Hobbies
  - Not many people know that I'm a competitive lockpicker/safecracker
    - You're safe – I don't do it professionally of course; I would say that wouldn't I ;)
  - You had me at rescue/retired greyhounds
  - Clearly enjoy food :) Although I'm not sure what I'm doing in front of a salad!!



# Trusted Research Environments Landscape View (research only – non-exhaustive)

## National TREs

- HSCNI - Honest Broker Service
- PHS - Scottish National Safe Haven
- NHS Digital
- UKSeRP
- ONS - SRS
- OpenSAFELY

## Commercial TREs

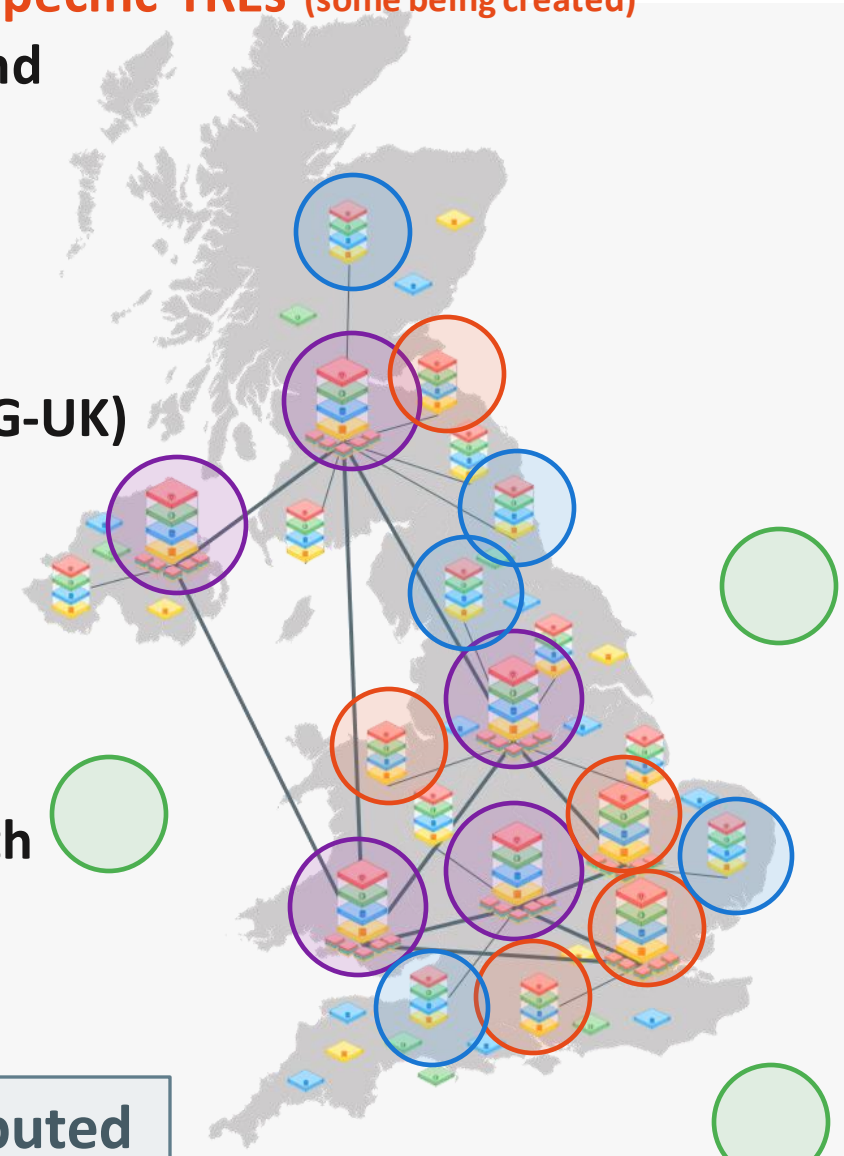
- AIMES
- AridhiaDRE
- AzureTRE
- AWS ServiceBench
- Lifebit
- ...

## International TREs

- Terra.bio
- TEHDAS
- MedCo
- ICODA
- ...

## Domain/Location-Specific TREs (some being created)

- Genomics England
- QResearch
- EMBL-EBI
- CPRD
- Turing Institute
- MRC-CLIMB (COG-UK)
- PIONEER
- eMedLab
- DISCOVER-NOW
- UKHSA
- CIPHA
- Our Future Health
- ...



The future of TREs is already here, just unevenly distributed

# What is a TRE?

A TRE is a **Trusted Research Environment**. Also known as 'Data Safe Havens', TREs are highly secure computing environments that provide remote access to health data for approved researchers to use in research that can save and improve lives.

## Why are they important?



**TREs make research safer.** Making data available through a TRE means that people can be **confident** that their personal health data is accessed **securely** and their **privacy protected**.

TREs help make **research efficient, collaborative** and **cost effective**, providing rich data that enables **deep insights** which will go on to improve healthcare and **save lives**.

TREs provide approved researchers with a **single location** to access valuable datasets. The data and analytical tools are all in **one place**, a bit like a **secure reference library**.

## How is my data safeguarded?

Health data should always be kept safe and secure, and used responsibly to ensure privacy. Health Data Research UK ensures these high standards are met by promoting the use of the 'Five Safes' model across all TREs.



### Safe People

Only trained and specifically accredited researchers can access the data



### Safe Projects

Data is only used for ethical, approved research with the potential for clear public benefit



### Safe Settings

Access to data is only possible using secure technology systems – the data never leaves the TRE



### Safe Data

Researchers only use data that have been de-identified to protect privacy



### Safe Outputs

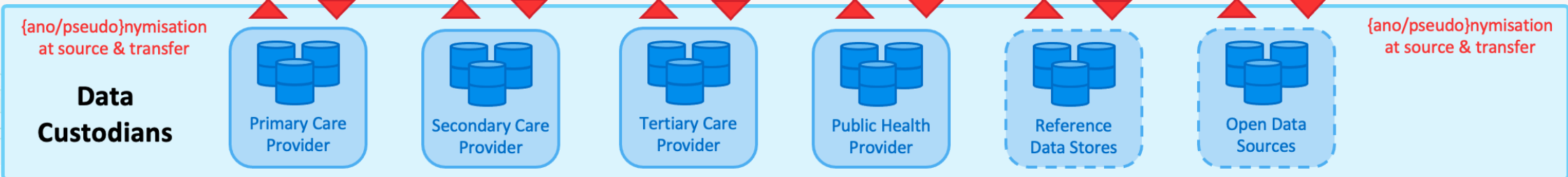
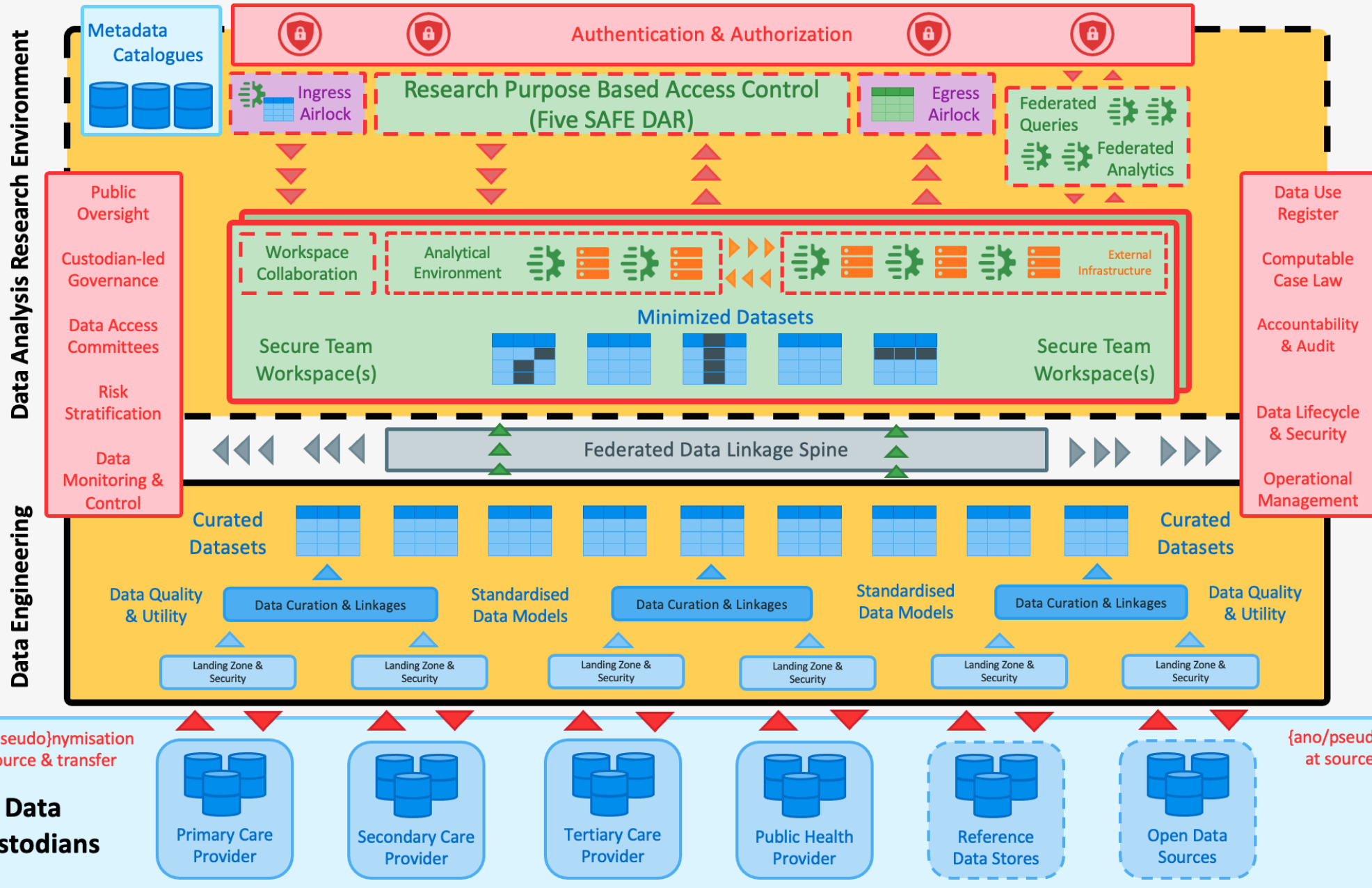
All research outputs are checked to ensure they cannot be used to identify subjects

Learn more about TREs and discover examples of how TREs are being used to enable life-saving health research.

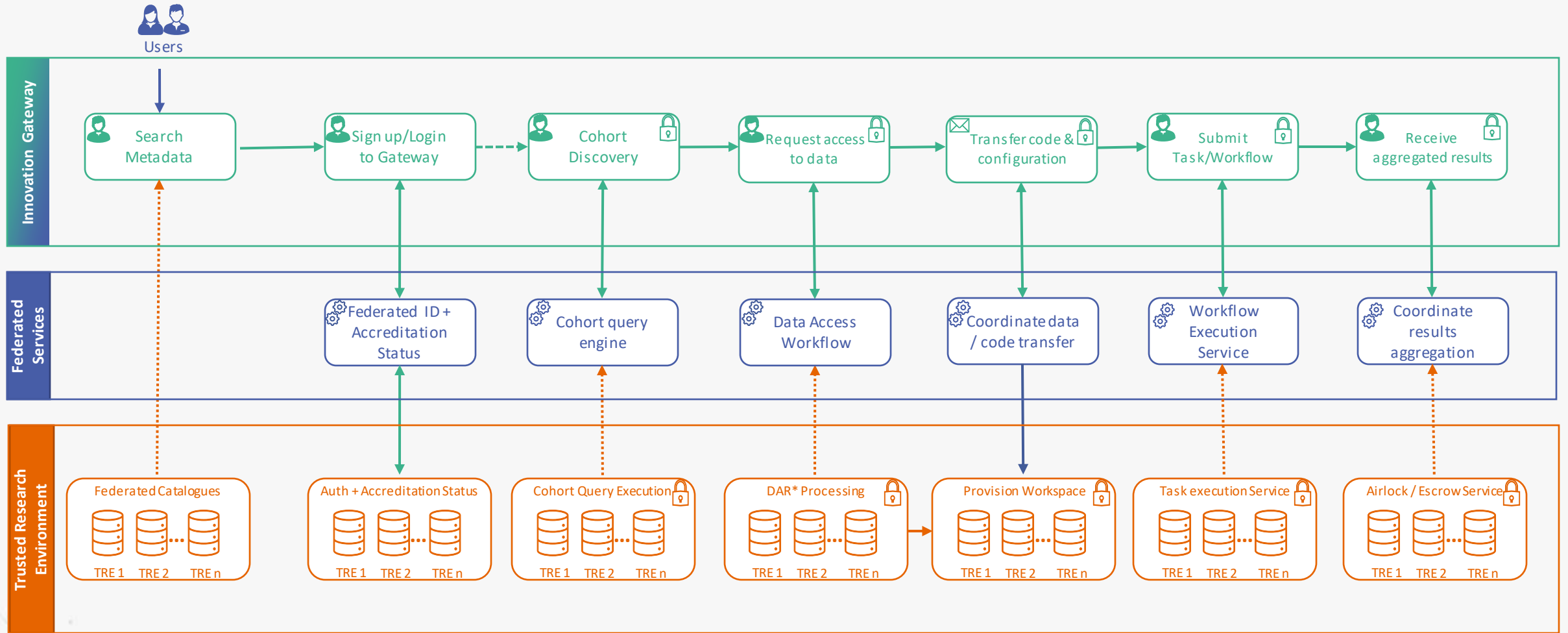
Learn more about TREs



# What is a Trusted Research Environment?

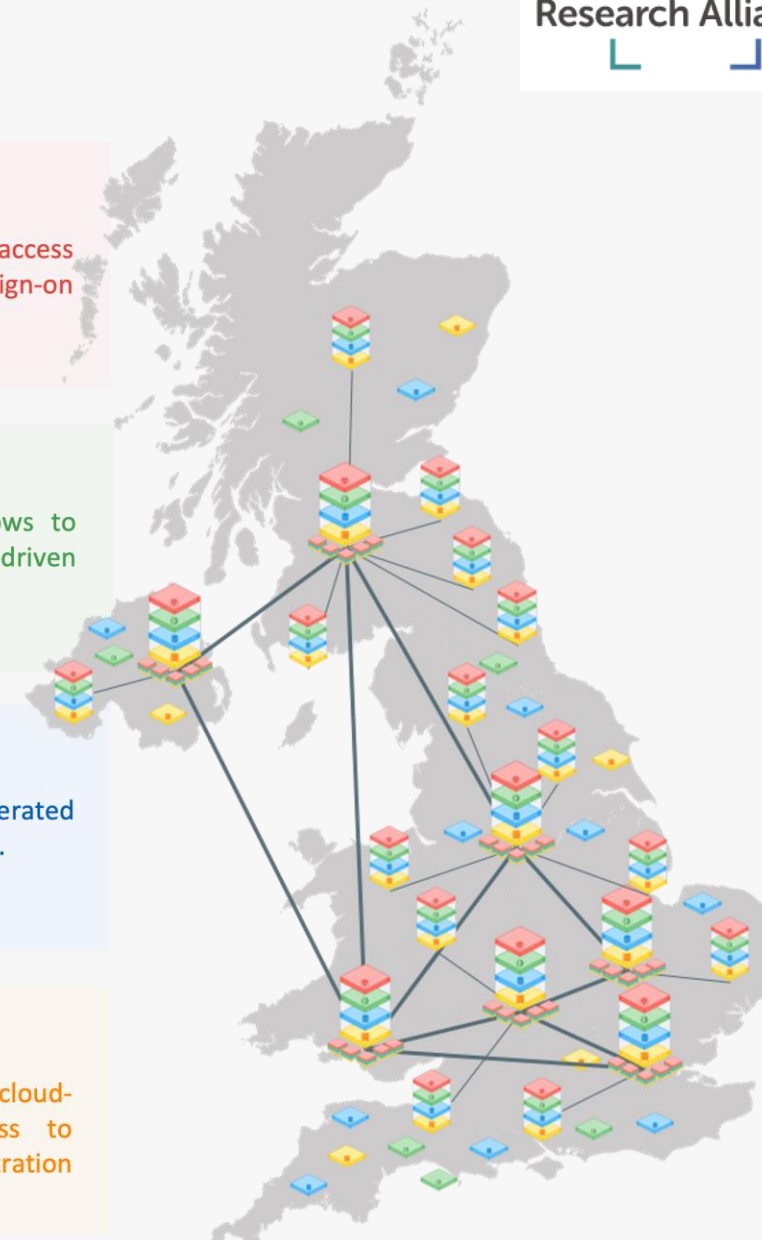
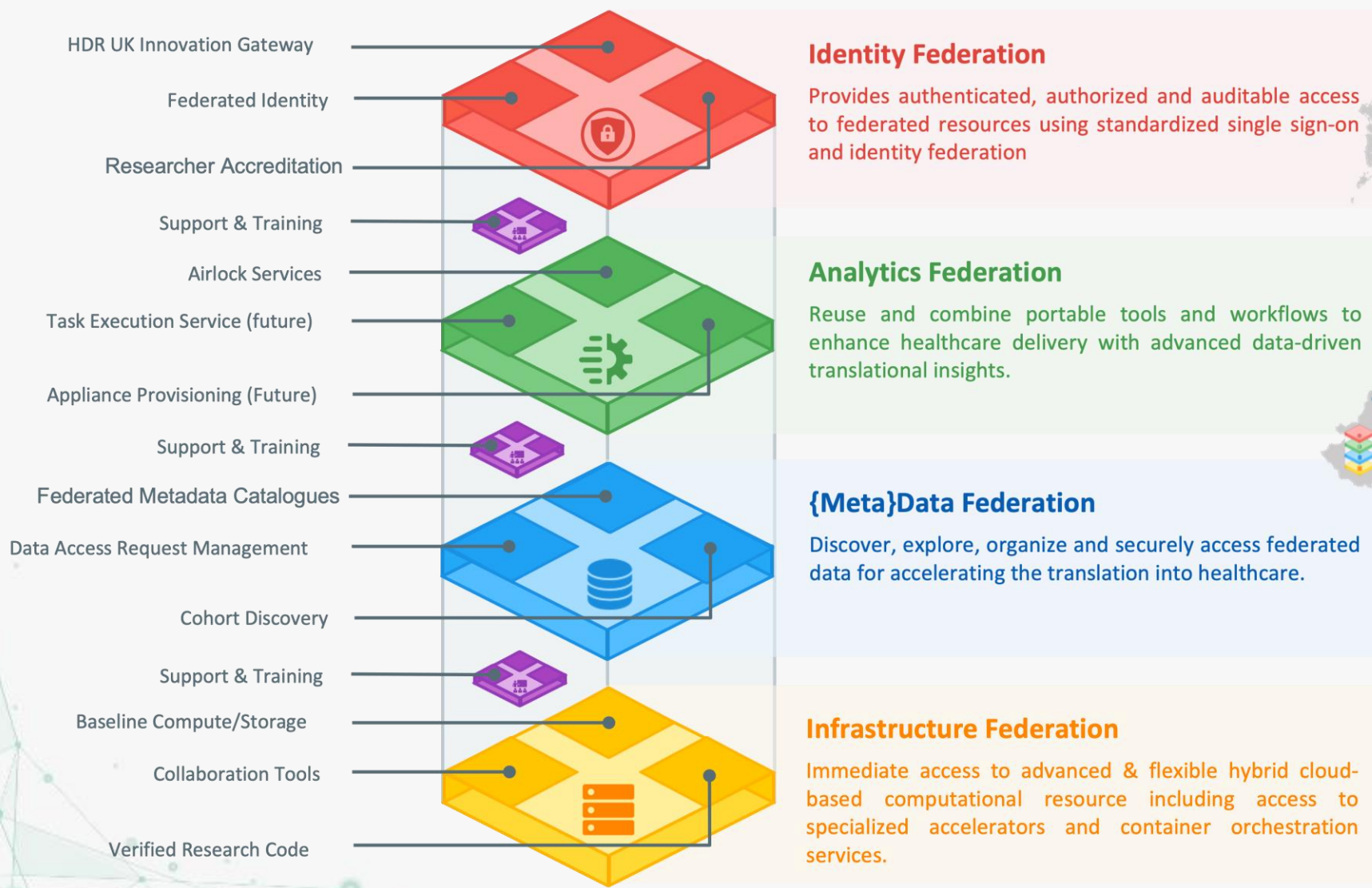


# Federated TRE User Journey – COVID-19 National Core Studies – Data & Connectivity



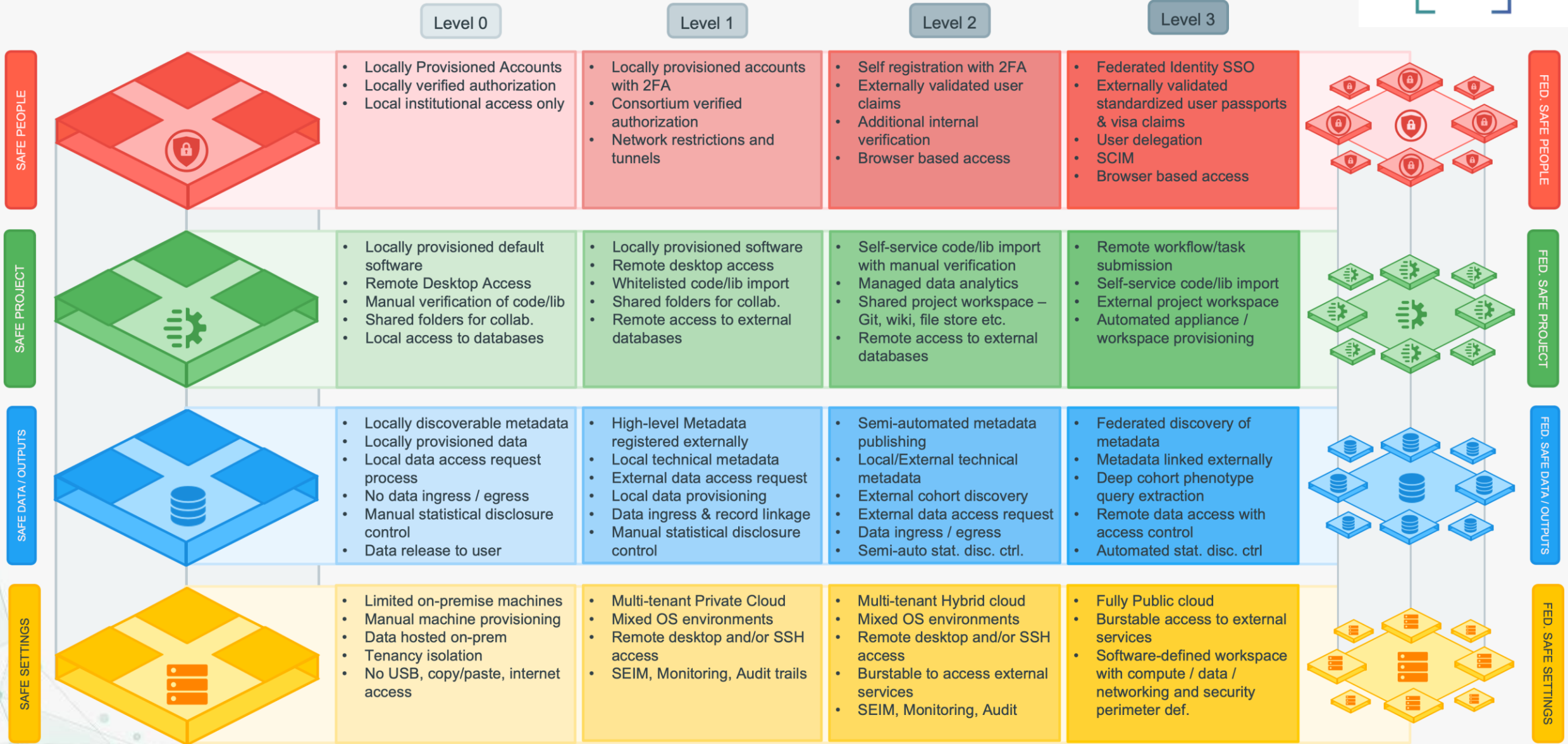
User action   
 Automated action   
 Process   
 Authenticated user action

# FOREST - Federated, Open & Resilient Ecosystem of Secure TREs



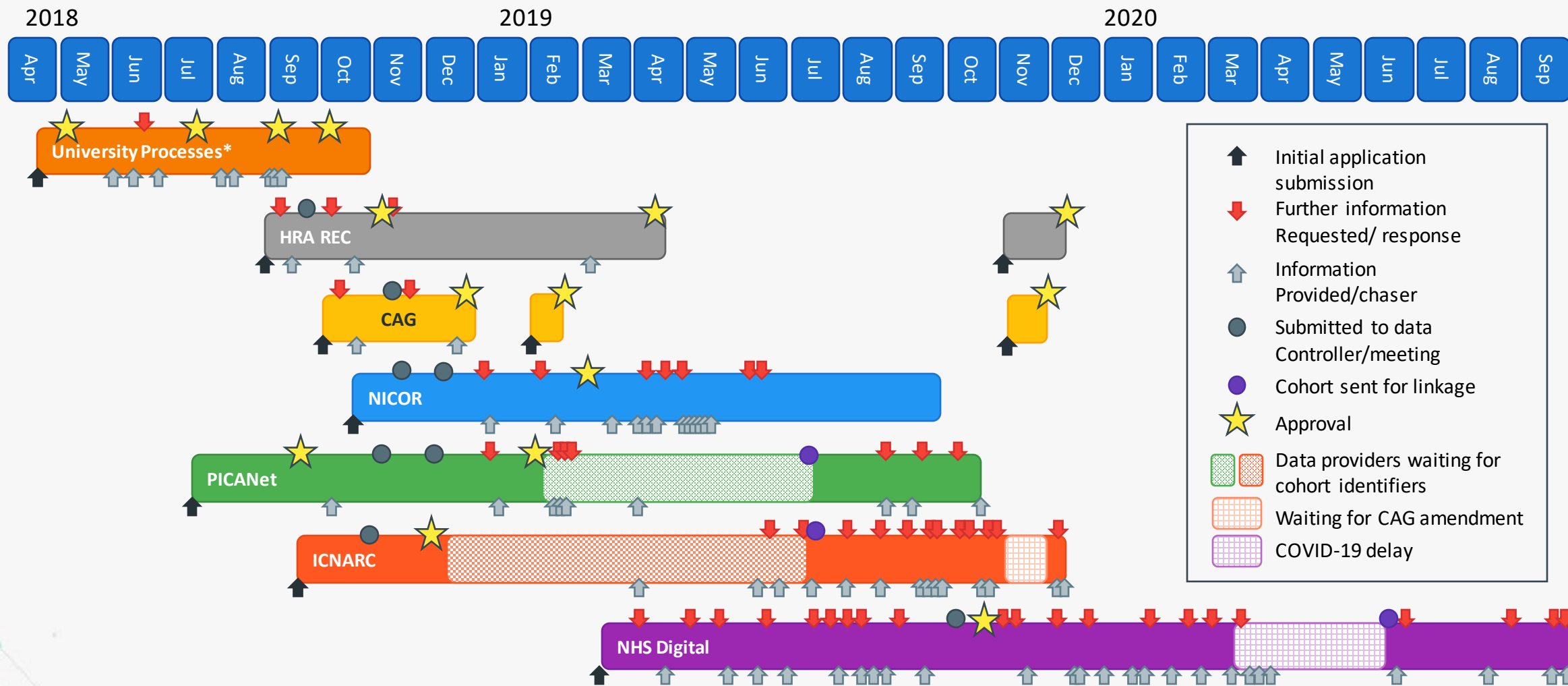
Layout for illustration purpose only

# FOREST – Federation Capability Maturity Model (Unbundling TREs)
















































# The road to hell is paved with good intentions...



# Data Governance Structures / Design Patterns

|   | Governance Structures                | Relationship | Data Availability | Permissions | Governance Challenges                 | IP, License & Contracts     |
|---|--------------------------------------|--------------|-------------------|-------------|---------------------------------------|-----------------------------|
|    | <b>Closed</b>                        | One → None   | Very Low          | Very Low    | Limits to collaboration               | <b>Enshrined in Law</b>     |
|    | <b>Restricted</b>                    | One → Some   | Very Low          | Low         | Limits to collaboration               | <b>Enshrined in Law</b>     |
|    | <b>One-to-one</b>                    | One → One    | Low               | Low         | Power asymmetry                       | <b>DSA, Contract</b>        |
|    | <b>Clique</b>                        | Some → Some  | Low/Medium        | Medium      | Trusteeship                           | <b>Consortium Agreement</b> |
|    | <b>Trusted Research Environments</b> | Some → Many  | High              | Low/Medium  | Trusteeship / user/project validation | <b>DSA, Contract</b>        |
|    | <b>Federated Analytics</b>           | Some → Many  | High              | Low/Medium  | Trusteeship / user/project validation | <b>DSA, Contract</b>        |
|   | <b>Federated Query</b>               | Some → Many  | High/Medium       | Low/Medium  | Trusteeship / user validation         | <b>DSA, Contract</b>        |
|  | <b>Open Access</b>                   | Some → Many  | High              | High        | Revocation of rights                  | <b>License</b>              |
|  | <b>Citizen Science</b>               | Many → Many  | High              | High        | Uneven capacity for analysis          | <b>Contract or License</b>  |

# Data Governance Design Pattern - Maturity Models (WIP)

| Governance Patterns     | Janeway  | Sisko  | Kirk  | Picard   |
|-------------------------|--|--|---|--|
| IP, License & Contracts |  Closed<br> Restricted     |  One-to-One   |  Clique<br> TREs<br> Fed. Analytics<br> Fed Query |  Open Access<br> Citizen Science     |
| Data Sensitivity        |  Closed<br> Restricted     |  One-to-One<br> Clique<br> TREs |  Fed. Analytics<br> Fed Query   |  Open Access<br> Citizen Science     |
| User Qualification      |  Closed<br> Restricted     |  Fed. Analytics<br> Fed Query  |  Clique<br> TREs  |  Open Access<br> Citizen Science     |
| Algorithmic Governance  |  Closed<br> Restricted |  Fed. Analytics<br> Clique   |  TREs<br> Fed. Query  |  Open Access<br> Citizen Science |

Not all governance structures and patterns are appropriate for all types of research collaborations

## Operationalizing Data Governance at scale – TREs are not a silver bullet

- Principles – FAIR, CARE?, SAFE?
- Requirements, Roles & Responsibilities, Rules of participation
- Interoperable Standards, Processes, Policies, Frameworks
- Modular Design Patterns, Software, Libraries
- Exemplars & Extensible Use Cases, Training, Awareness, Community Engagement



We need a set of **PRIME Directives** to help operationalize, interoperate and federate across data governance patterns

## The future - building on this excellent foundation and partnerships to deliver a step change in benefits for UK science and population

- **Data as Infrastructure** – building blocks to create interoperable global networks focused on collaboration and sharing
- **Clinical/Scientific Driver Programmes** that help guide our development efforts and pilot tools
- **UK wide** and at the centre of an international collaborative network of science
- **Diverse data types beyond NHS data** (omics, biomedical, wearables, social sciences)
- **Thought leadership and expertise for innovative approaches to trust, governance and standards** for data collection and access
- Exemplar of **team science** and **partnership** working
- **Patient Involvement and Engagement** at all stages of the research lifecycle
- **Training the current and next generation of scientists** to responsible, ethical and equitable use of AI

**HDRUK**  
Health Data Research UK

UK Health Data  
Research Alliance

## Thank you for listening

Here's some obligatory holiday snaps from the peak district

[bit.ly/HDRUK-TRE](https://bit.ly/HDRUK-TRE)



[hdruk.ac.uk](https://hdruk.ac.uk)



[healthdatagateway.org](https://healthdatagateway.org)



Email: [susheel.varma@hdruk.ac.uk](mailto:susheel.varma@hdruk.ac.uk)

Twitter: [@susheelvarma](https://twitter.com/susheelvarma)

Do we want a “FORESTRE Commission” responsible for increasing the value of TREs – expanding, promoting, protecting the sustainable management of TREs across all four nations?