The Data Domains: <u>The Present Picture</u>:
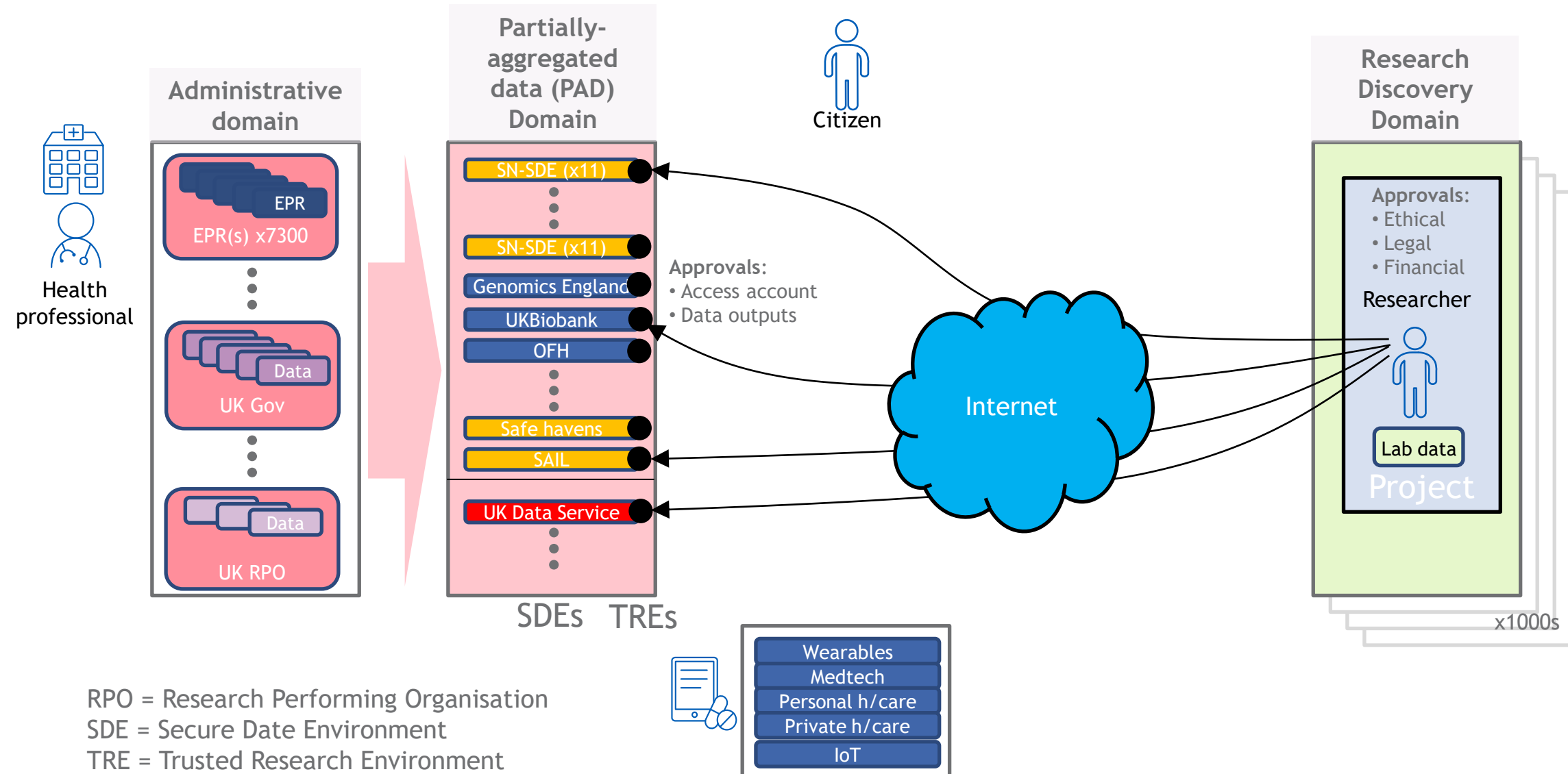
* Datasets have add-on TREs.

* Approved projects negotiate (again) for approved access. Data chain of custody fragmented.

* Duplication of admin, extending time to research output.

RPO = Research Performing Organisation
SDE = Secure Date Environment
TRE = Trusted Research Environment

The Data Domains: <u>The Goal</u>:

* Peer Connected SDEs to project TRE platforms. Orchestrated project configuration
* Open logs and analytics for infringement of data policies and risks.
* SHARE architecture – extending the control of SDEs into the collaboration projects.
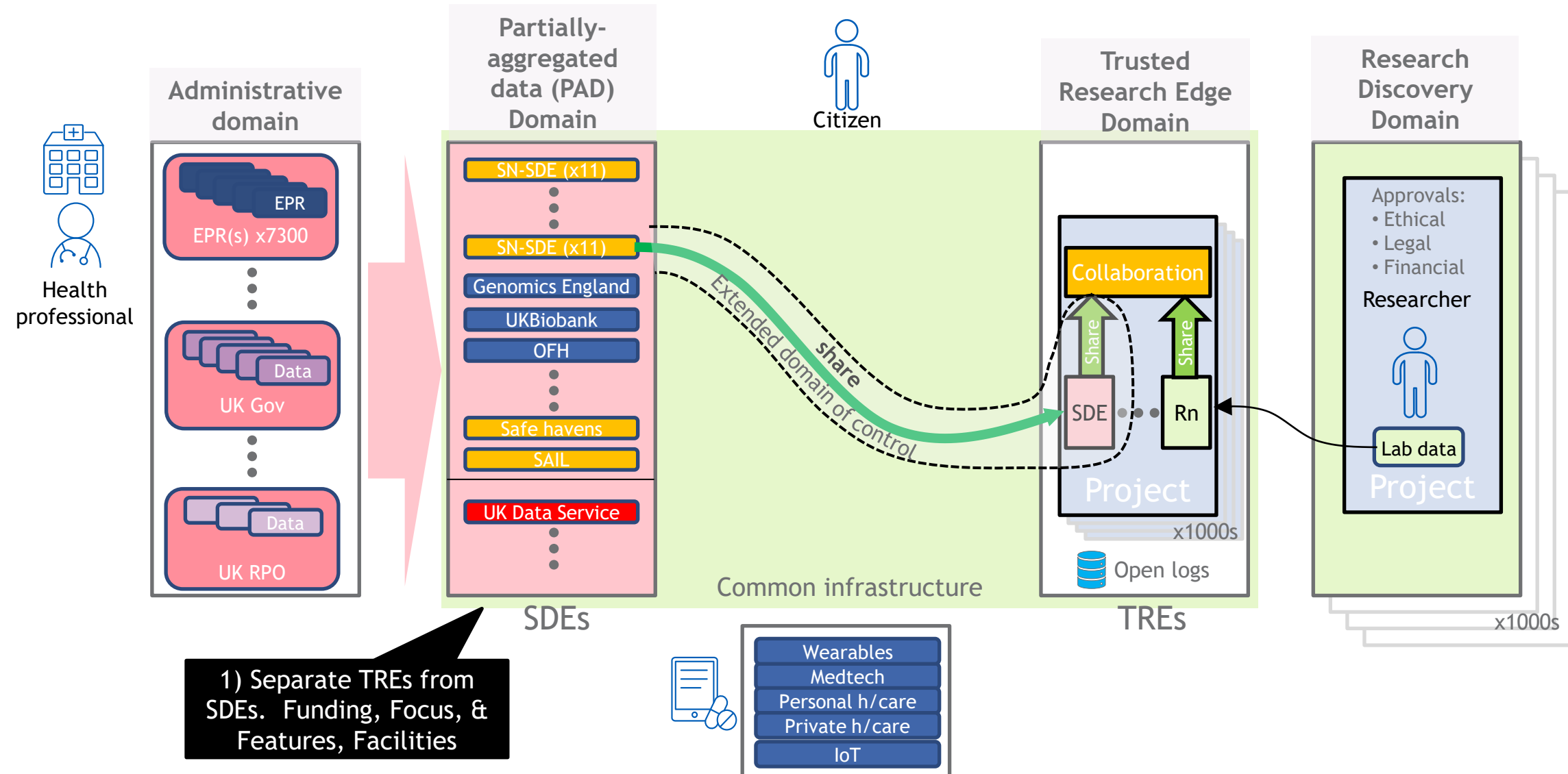
So emerging Definitions... needing a good "scrubbing"

See Crick TRE white paper: https://doi.org/10.25418/crick.23626653.v1

Object	Acronym	AKA	Definition
Secure Data Environment	SDE	Data Safe Haven	• A master repository of datasets. • It curates data and has established ELT / ETL routines that collect this data from respected, definitive sources. • It provides Info Governance to manage access from accredited TREPs as platform peer-peer partners. • It makes views of its data for projects and shares these to the "projects via control of its TREP owned space in each project TRE. "Safe data"
Trusted Research Environment Platform	TREP		• This is a platform that has facilities to set up, manage and archive Project Trusted Research Environments (PTREs) and provides the "settings" for projects to operate. • Research performing organisations become accredited to the TREP to operate their projects. "Safe projects", "Safe people" • It provides the facilities to add users into predetermined and defined roles within the Project TREs. • It provides compute and storage resource, and monitors usage to allow usage /spend control and accounting / billing. • Audit and Compliance reporting to enable projects to discharge their obligations.
Project TRE	PTRE		• The project TRE is specific to an individual Project and provides a dedicated controlled compute and storage facility for all the partners in the project. • Each collaborating org in the project has their own private space connected to the collaboration / research project space. • This enables people to work on projects to complete the research objectives within the ethical / legal and financial constraints / agreements. • The project and project TRE has accountability for "safe outputs".
Project			• ...
View			• ...

An architecture for extending the domain of control - custody of data.

Axiom: Data chain-of-custody is a corollary of consent management
## How can this fit with SDE / TRE infrastructures?

Pete Barnsley, PhD
Head of Special Projects, ITO, The Francis Crick Institute

- What technical architectures help **govern research data access** in a world with interconnected SDEs and TREs?
- How does **the citizen's "consent" to** share get implemented?

- How does the **SDE's domain of control** extend and connect to the project?

- What architectures offer this capability?
- How could the SDE and TRE interconnect?
- What are the pros and cons of each interconnect option?
- What would we propose should be built for the UK?

Collaboration opportunities: points 1) to 6):
* Bring the data to the project.
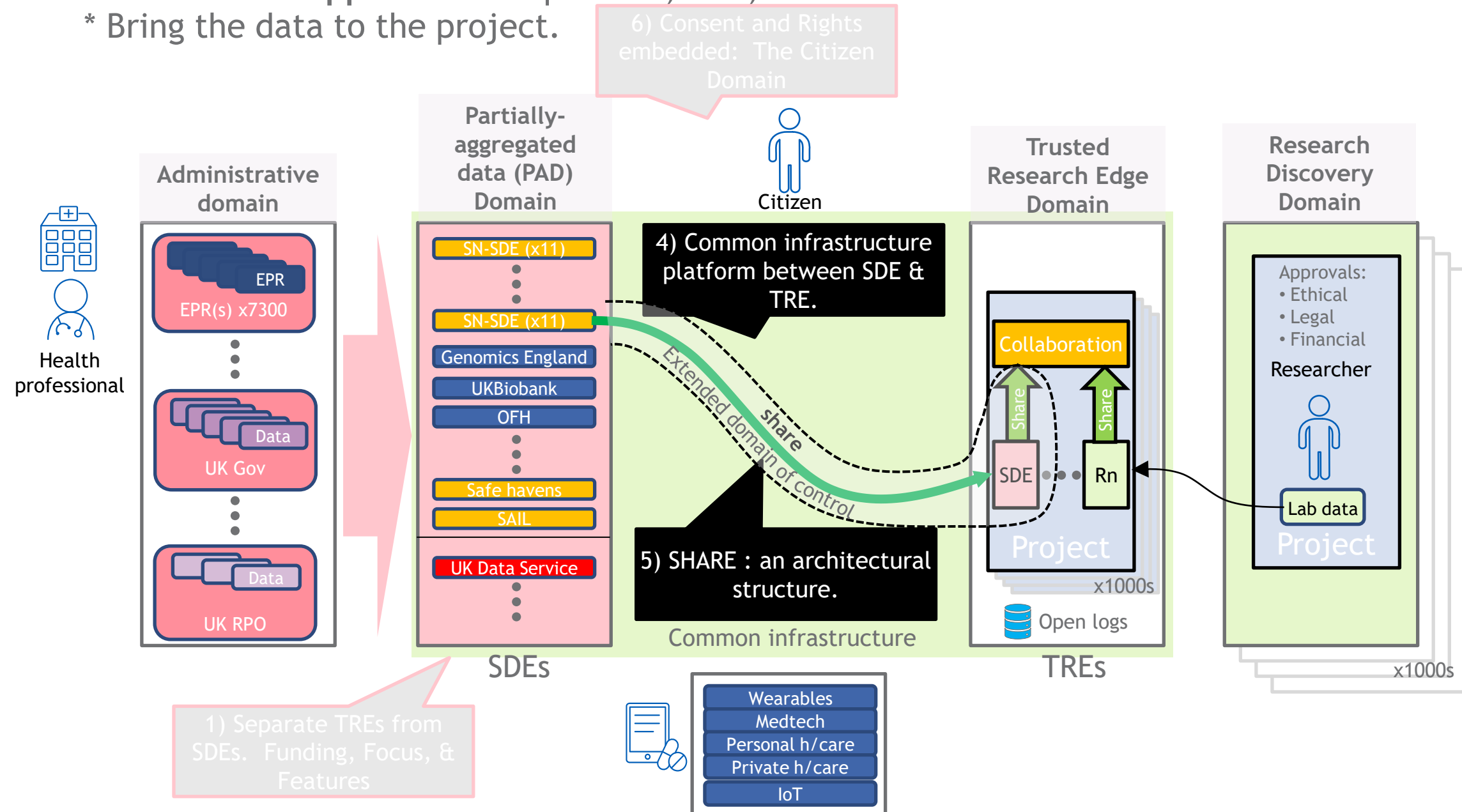
An architecture for extending the domain of control - custody of data

Lead:

Pete Barnsley   (Head of Special Projects, Francis Crick Institute, London)

1. **Set up a first open session in January**
   (to allow diaries **proposal is 10-11am Wednesday 10th and 2-3pm Tuesday 16th**)
   - Expectations of people within the workstream
   - Ideas and scope.
   - Plan for completing output.
   - Liaising with other related workstreams – e.g. Federation

2. **Objective 1: Plan to present at the March TRE Community meeting**
   - Interim progress report.

3. **Objective 2: White paper**
   - Recommendations for retaining "domain of control" and a "chain of custody" for data.
   - Implications for national initiatives and on other TRE Community workstreams, e.g. federation.