



# Federated Architecture Blueprint - Part 5 (Delivery Plan)

DARE UK Delivery Team


Version 2.2 final

November 2024



**FOR CONSULTATION & COMMENT**

**Licence**

This work © 2024 by HDR UK and other members of the DARE UK consortium is licensed under CC BY-NC-SA 4.0 The Creative Commons license icons, including the CC logo, a person icon (BY), a crossed-out dollar sign (NC), and a circular arrow (SA).

## FOR CONSULTATION & COMMENT

### Document control

Version	Date	Authors/Reviewers	Notes
0.6	22/03/2023	Rob Baxter	First complete draft.
0.7	31/03/2023	Fergus McDonald, Hans-Erik Aronson	DARE UK internal review.
1.0 initial	13/04/2023	Rob Baxter	For publication and public comment.
1.1	03/08/2023	Rob Baxter	Updated. Feedback until end June 2023 incorporated.
1.2	11/08/2023	Rob Baxter	Version for internal review.
1.3	15/08/2023	Fergus McDonald, Emily Jefferson	DARE UK & HDR-UK internal review.
1.4	25/08/2023	Rob Baxter	Updated. Greatly expanded Executive Summary. Version for internal review.
1.5	04/10/2023	Fergus McDonald, Emily Jefferson	DARE UK & HDR-UK internal review.
1.6 interim	18/10/2023	Rob Baxter	For broader circulation and comment.
2.0 draft	11/12/2023	Rob Baxter	Incorporated revisions and lessons learned from Driver Projects and wider engagements.
2.0A draft	12/12/2023	Rob Baxter	Incorporated review feedback from SACRO project PI.
2.0B draft	08/01/2024	Rob Baxter	Incorporated review feedback from TRE-FX project PIs.
2.0C draft	29/02/2024	Fergus McDonald, Emily Jefferson	DARE UK & HDR-UK internal review.
2.0D draft	28/03/2024	Rob Baxter	Final tidy-up, incorporating research use-cases from February 2024 workshop.
2.0E draft	13/06/2024	Fergus McDonald, Emily Jefferson, Caole Goble, Phil Quinlan, Simon Thompson	Partner review.
2.0F draft	05/08/2024	Rob Baxter, Heikki Lehväslaiho	Fixed error in Chapter 8, prototype descriptions.
2.1 draft	30/08/2024	Rob Baxter	Restructuring across Chapters 2-4; realignment and rationalisation of user roles.
2.1 Part 1..5	19/09/2024	Rob Baxter	Separation into multiple parts for release.
2.2	31/10 2024	Emily Jefferson	DARE UK & HDR UK internal review.
2.2 final	11/11 2024	DARE UK	For release.

**FOR CONSULTATION & COMMENT**

## Contents

---

Document control .....	3
Contents.....	4
About document versions .....	5
1. Development and delivery approach.....	6
2. Prototyping and technology selection.....	7
2.1. Core services: technology evaluation.....	7
2.2. Interfaces and other services: community driver projects.....	7
3. Technology proof-of-concept .....	8
3.1. Scenario 1: basic data exchange.....	8
3.2. Scenario 2: linked data exchange .....	8
3.3. Scenario 3a: remote direct query (single) .....	8
3.4. Scenario 3b: remote direct query (federated).....	9
3.5. Scenario 4a: remote indirect query (single).....	9
3.6. Scenario 4b: remote indirect query (federated) .....	9
4. Minimal viable product .....	11
4.1. Test and validation .....	11
4.2. Evolution .....	11

**FOR CONSULTATION & COMMENT**

## About document versions

---

This document is Part 5 (Delivery Plan) of the *Federated Architecture Blueprint* for DARE UK. It defines a potential approach for an overall architecture for a network of sensitive data sources and secure analytical services in terms which are broadly—and deliberately—**technology neutral**. Choices of implementation technology are not dealt with here, nor are details of costs, benefits and delivery plan.

This document covers architecture version 2. It refines the model of a federated network infrastructure from the “initial” and “interim” versions, builds further on the “data layer” and most significantly draws in lessons and learnings from the 2023 DARE UK Driver Project programme.

**FOR CONSULTATION & COMMENT**

## **1. Development and delivery approach**

---

Our adopted design philosophy favours an incremental approach to delivering the federated network architecture: introducing (and enforcing) a common low-level foundation while aiming for minimal disruption to existing services and supporting maximum innovation at application level. This short document sketches a phased delivery approach to building the first operational version of a federated TRE and data network. Note that we do not cover the long-term operational funding model of such a federated network of TREs here, although we do observe that such a model is absolutely critical for this network to succeed long-term in delivering better research outcomes for the UK.

**FOR CONSULTATION & COMMENT**

## **2. Prototyping and technology selection**

---

Suitable technologies to deliver the Federation core services should first be explored and selected. Two different approaches can be used, depending on the technology readiness level (TRL) required<sup>1</sup>.

### **2.1. Core services: technology evaluation**

Core services provide the secure, trustworthy backbone of the entire Federation. These should be selected from existing solutions, proven in operation (i.e., TRL 9).

We recommend convening a community-wide panel to draw up a shortlist of potential solutions and then commissioning a series of evaluation projects against a common “proof-of-concept” brief.

### **2.2. Interfaces and other services: community driver projects**

Securing the foundation layer allows for greater innovation at the interface and application level without increasing risk. The core interface services that run on top of the data exchange foundation can thus be drawn from a wider ecosystem. Experimentation between Participants is possible at this level without undermining the security of data exchange.

We recommend commissioning research and development projects to investigate different technological approaches to the required core services. DARE UK’s Phase 1b Driver Projects (2023) are a model approach<sup>2</sup>.

---

<sup>1</sup> Technology Readiness Levels. See [https://en.wikipedia.org/wiki/Technology\\_readiness\\_level](https://en.wikipedia.org/wiki/Technology_readiness_level)

<sup>2</sup> See <https://dareuk.org.uk/our-work/phase-1-driver-projects/>.

**FOR CONSULTATION & COMMENT**

### 3. Technology proof-of-concept

---

Using selected technologies, a proof-of-concept (PoC) system can be deployed against a number of test scenarios. Note that functionality and correct operation can be tested in all these scenarios without the need for any sensitive data.

Scenarios 1 and 2 below cover “traditional” TRE operation where data are moved into a secure environment for analysis. Scenarios 3 and 4 develop the newer remote-query model.

Note that all these scenarios are technical proofs-of-concept that demonstrate the required functionality of foundational and core components. They do not address data interoperability or information governance.

#### 3.1. Scenario 1: basic data exchange

This is the base scenario involving the core Federation Services and secure data exchange between two TREs, one acting as a data provider and one as an analytical service.

Required components:

- 1 x Federation Services (Core);
- 1 x TRE: Security Server (Core); SDZ (interfaces: Data Ingress);
- 1 x TRE: Security Server (Core); SDZ (interfaces: Data Egress).

Required concepts:

- Identities: Participant; Project; Dataset; Data Extract;
- Structured Data Objects: Data Extract.

#### 3.2. Scenario 2: linked data exchange

This scenario extends the first with an additional data provider TRE and introduces an Index Service.

Required components:

- 1 x Federation Services (Core);
- 1 x TRE: Security Server (Core); SDZ (interfaces: Data Ingress, Index);
- 2 x TRE: Security Server (Core); SDZ (interfaces: Data Egress, Index);
- 1 x Index Service: Security Server (Core); interfaces: Index.

Required concepts:

- Identities: Participant; Project; Dataset; Data Extract; Linkage Spine;
- Structured Data Objects: Data Extract; Linkage Spine.

#### 3.3. Scenario 3a: remote direct query (single)

This scenario exercises the movement of direct queries rather than the movement of data and can be viewed as complementary to Scenario 1. Recall that “direct queries” are fully encapsulated within Query data objects.

Required components:



## FOR CONSULTATION & COMMENT

- 1 x Federation Services (Core);
- 1 x TRE: Security Server (Core); RAZ (interfaces: Query (direct), Response);
- 1 x TRE: Security Server (Core); SDZ; QMZ (interfaces: Query (direct), Response).

Required concepts:

- Identities: Participant; Project; Dataset;
- Structured Data Objects: Query; Response (query).

### 3.4. Scenario 3b: remote direct query (federated)

This scenario extends Scenario 3a to include a second data provider and tests the splitting of a query to run against each independently. Note that this requires more sophisticated data presentation and query handling than Scenario 3a.

Required components:

- 1 x Federation Services (Core);
- 1 x TRE: Security Server (Core); RAZ (interfaces: Query (direct), Response);
- 2 x TRE: Security Server (Core); SDZ; QMZ (interfaces: Query (direct), Response).

Required concepts:

- Identities: Participant; Project; Dataset;
- Structured Data Objects: Query; Response (query).

### 3.5. Scenario 4a: remote indirect query (single)

This scenario exercises the movement of indirect queries rather than the movement of data. Recall that “indirect queries” are **not** fully encapsulated within Job Request data objects but instead refer to (or “point to”) an analysis workload hosted on a third-party Software Service which must be retrieved by the participating TREs prior to execution.

Required components:

- 1 x Federation Services (Core);
- 1 x TRE: Security Server (Core); RAZ (interfaces: Query (indirect), Response);
- 1 x TRE: Security Server (Core); SDZ; QMZ (interfaces: Query (indirect), Response; Software);
- 1 x Software Service (research artifacts): Security Server (Core); interfaces: Software.

Required concepts:

- Identities: Participant; Project; Dataset;
- Structured Data Objects: Job Request; Response (job); Job Payload Artifact.

### 3.6. Scenario 4b: remote indirect query (federated)

This scenario exercises the movement of indirect queries rather than the movement of data.

This scenario extends Scenario 4a to include a second data provider and tests the splitting of a query to run against each independently. Note that this requires more sophisticated job and query handling than Scenario 4a.

## FOR CONSULTATION & COMMENT

### Required components:

- 1 x Federation Services (Core);
- 1 x TRE: Security Server (Core); RAZ (interfaces: Query (indirect), Response);
- 2 x TRE: Security Server (Core); SDZ; QMZ (interfaces: Query (indirect), Response; Software);
- 1 x Software Service (research artifacts): Security Server (Core); interfaces: Software.

### Required concepts:

- Identities: Participant; Project; Dataset;
- Structured Data Objects: Job Request; Response (job); Job Payload Artifact.

**FOR CONSULTATION & COMMENT**

## 4. Minimal viable product

---

A successful technology proof-of-concept for (at least) scenarios 1 and 2 should be developed into a minimal viable product (MVP). Scenarios 3 and 4 (and other functionality) can be introduced later through evolution and improvement.

**Note that MVP development here is not principally a technical activity.** The journey from proof-of-concept to MVP should focus on developing the required governance framework around data exchange, linkage and project identities.

The end product of this phase is a limited deployment of a federated TRE network suitable for use with real data.

### 4.1. Test and validation

Alongside the development of an MVP a test and validation approach should be developed. This should involve the deployment of a mirror version of the PoC system and the instigation of a dedicated adversarial test team (a “red team” in security engineering jargon<sup>3</sup>).

We recommend including a dedicated red team testing component in future operational plans for the SDRI Federation.

### 4.2. Evolution

Once in place the MVP can be expanded and extended incrementally in scope and functionality:

- Scope: new TREs and other services can be added to the network by deploying Security Servers and supporting appropriate interface Services;
- Functionality: new interface Services can be developed and incorporated into the Federation’s “working set” as technology evolves.

In both cases, how changes are made and approved are key decisions required of Federation governance.

---

<sup>3</sup> See [https://csrc.nist.gov/glossary/term/red\\_team](https://csrc.nist.gov/glossary/term/red_team) for a definition of “red team”. The NCSC also has a good discussion of red-teaming in machine-learning system design at <https://www.ncsc.gov.uk/collection/machine-learning/requirements-and-development/design-for-security>