

# DARE



## SUMMARY REPORT: Recommendations for TRE governance to support innovation and interdisciplinary

## **SUMMARY REPORT: Recommendations for TRE governance to support innovation and interdisciplinary research**

### **Purpose**

- 1. Dare outcome – money well spent**
- 2. Public outcome – ditto**
- 3. Summary of recommendations for TRE (to accompany fuller governance framework doc)**

### **Contents**

## 1. Project team

---

This project is partly funded by UK Research and Innovation Grant Number MC\_PC\_21029 as part of Phase 1 of the DARE UK (Data and Analytics Research Environments UK) programme, delivered in partnership with Health Data Research UK (HDR UK) and ADR UK (Administrative Data Research UK).

The project has been led by the DataLoch team at the University of Edinburgh, in collaboration with Public Health Scotland's (PHS) Electronic Data Research and Innovation Service (eDRIS). DataLoch is a data service that curates health and social care data for 1.6m people in South-East Scotland. Approved applicants receive de-identified data extracts via the National Safe Haven – a Trusted Research Environment (TRE) managed by eDRIS.

**Professor Nicholas Mills:** Professor Mills is a British Heart Foundation Professor of Cardiology at the University of Edinburgh and Consultant Cardiologist at the Royal Infirmary of Edinburgh. He is the Senior Responsible Officer and Academic Lead for a £68 million Data Driven Innovation in Health & Social Care Programme funded by the Edinburgh & South-East Scotland City Region Deal.

**Atul Anand, DataLoch Clinical Lead:** Atul is an NHS Consultant Geriatrician and Senior Clinical Research Fellow at the University of Edinburgh. His role within DataLoch and in this project has been working with the technical team and clinical colleagues to bring a clinical perspective to the design.

**Kathy Harrison, DataLoch Programme Lead:** Kathy is a project management professional responsible for launching the DataLoch data service, including the service's robust, innovative, and agile governance process. She is the overall project manager for this project.

**Amy Tilbrook, Data Ethics Professional and DataLoch Information Governance:** Amy is an accredited ethics professional, with extensive experience of health data research governance. She has been responsible for leading the governance definition work package (WP1).

**Stuart Dunbar, PPIE lead:** Stuart has a PhD in science engagement and has been responsible for engaging with the public to understand perspectives on non-traditional researcher access to data (WP3). These perspectives have been key to the recommendations on governance conditions.

**Carole Morris, eDRIS Head of Service:** The PHS eDRIS team have run the National Safe Haven for many years, supporting the administration of hundreds of research projects. Carole has been responsible within this project for consulting on the national requirements from the TRE prototype.

**Michelle Evans, Talent Programme Director:** Michelle is the Talent Lead for the DDI Health and Social Care programme at the Usher Institute and Director of the online MSc in Data Science for Health and Social Care, responsible for the shaping the future of a major new health & social care programme for Edinburgh & South East Scotland as part of the DDI initiative at the University of Edinburgh. Michelle has been responsible for leading the work package to develop training content and make it available on a public learning platform (WP2).

Talent team members: Chris, Richard: Responsible for designing the training module (WP2), and developing the module delivery and technical solution.

## 2. Project summary (Max 350 words)

---

The team have explored the barriers to the use of TREs by non-traditional researchers from different disciplines and investigated public perspectives around non-traditional user access to health and social care data. Non-traditional users are those researchers outside academia, or innovators from different disciplines and

organisations, including software developers, AI experts and others from third- and private sector organisations. These organisations are critical to driving improvements to health systems and outcomes across the United Kingdom. While TREs support privacy by design, access by non-traditional users has been limited and remains largely aspirational.

This project supports DARE UK programme objectives to enable UK researchers and innovators from a mix of disciplines to harness sensitive datasets to deliver research and innovation in the public good.

This report outlines recommendations for TRE governance for non-traditional users. This is based on 3 workstreams:

1. Design of a governance model prototype and governance standards for access to sensitive data that adapts the successful Five Safes framework for non-traditional researchers.
2. Delivery of an online learning and accreditation module for non-traditional researchers working with health and social care data, to ensure robust standards are clearly achieved before granting data access.
3. Public engagement to understand perceptions on what is expected from 'non-traditional' TRE users – such as those from the third and private sectors – to be considered trustworthy and credible and what public concerns exist about their use of TREs.

## 2.1. Workstream 1: Governance Model

TREs are widely used to support secure data access for researchers, enabling collaborative analysis without compromising data privacy. DataLoch's approved applicants receive de-identified data extracts via the National Safe Haven – a TRE managed by eDRIS. However, this TRE was designed primarily for clinical and academic researchers using traditional statistical software. This project has identified a number of specific recommendations for non-traditional access of TREs and has used these to create a governance framework for DataLoch in Scotland.

## 2.2. Workstream 2: Training Module

Existing user training requirements aimed at traditional researchers, such as MRC courses on research, GDPR and confidentiality, reflect the needs of most traditional TRE users. However, non-traditional users particularly those not from research backgrounds may need an additional grounding in health and care data challenges and their responsibility in accessing sensitive data via a TRE. Through the creation of a new online professional development learning module targeted at the non-traditional user, this project aims to support these users, to ensure robust standards are clearly achieved before granting data access.

The interactive learning module is hosted on the globally renowned Moodle virtual learning environment (VLE) platform. Users will create their own account using email based self-registration, stating which organisation they are from. Moodle provides reporting features that can be filtered by user details, with built in evaluation and analytics.

It is designed to be a standalone, self-directed module and covers governance and standards for access to sensitive health and social care data that adapts the Five Safes framework. By design it can be integrated into existing training for safe researchers, supporting TRE users, current training providers, and data controllers across the DARE UK community.

## 2.3. Workstream 3: Public Engagement

Public perspectives about data access to support traditional researchers have been explored in previous engagement exercises. However, there is less understanding of public perspectives on the use of TREs by non-

traditional researchers – such as software developers, Artificial Intelligence experts, and representatives from third- and private-sector organisations.

This project completed public consultation – conducted through Ipsos Scotland – consisting of two elements. First, a survey of 595 South-East Scotland residents sought top-level perspectives on which organisations should have access to health data, and the conditions under which this access could be permitted.

Second, two deliberative workshops involving approximately 40 local residents (broadly representative of the local population in Scotland) sought insights that would inform the design of DataLoch’s governance framework. Participants were posed two overarching questions:

- What are the benefits and risks of health data being used by non-traditional researchers?
- What needs to be in place for access to health data by non-traditional researchers to be considered trustworthy and acceptable?

The sessions were broadly divided into two phases: the learning phase (session one), where participants got to know about the key concepts of importance, and the deliberation stage (session two) where they discussed the issues in detail and developed their conclusions.

## Key findings

Participants were generally positive about the benefits of health data research and the use of TREs, although there was concern about potential data breaches and the possible consequences. The principles and red lines developed by workshop participants centred on the following themes:

- Commercial interests** – while it was acknowledged that research by private-sector companies may have public benefit, such as the development of medicines and improvements in treatment, there was concern about data being accessed by organisations where their sole motivation was to make profit.
- Transparency of purpose** – to foster trustworthiness, it was felt that every researcher – either traditional or non-traditional – should be required to provide a clear justification for why they need data, what data they need and how they will use it.
- Health-related impacts** – related to the principles above, there was a desire for non-traditional researchers to be able to demonstrate the ultimate public benefits of their proposed research and that any new innovation will show measurable benefits over traditional approaches.
- Ethics panel** – these were seen as an effective way of providing scrutiny of data access proposals. Panels should involve health care professionals, experts in health data research and the ethical considerations of data-driven innovation, and members of the public.
- Terms and conditions** – it was assumed that traditional researchers were already subject to codes of conduct and that non-traditional researchers should be held to a similar standard.
- Monitoring and review** – having a mechanism to check if researchers have used data in alignment with their original stated purposes was seen as a way of providing additional reassurance for the process.

The public workshops suggest the public recognise the balance needed between protecting patients’ data, and making sure access is possible for research that brought innovation and improvement in health services. Participants acknowledged the potential benefits of non-traditional researchers accessing health data. However, this did not mean there was blanket acceptance of non-traditional researchers accessing health data. Some types of non-traditional researcher, particularly those involving organisations operating outside of the health care field and those operating for profit, raised particular concern.

## 3. TRE Governance recommendations

---

The recommendations are articulated within the Five Safes framework which was originally developed by the [Office for National Statistics \(ONS\)](#) and is now used by research data centres worldwide. The recommendations are intended as guidance for TREs looking to extend their researcher safe access governance to non-traditional users including private- and third-sector.

The recommendations assume that baseline governance controls are in place that reflect the Five Safes high standards for research, and are focussed on additional controls or augmentation of existing controls required specifically to support non-traditional user access to data via these TREs.

### 3.1. Safe Projects

**3.1.1. Any use of personal data must have an explicit aim to deliver public value. In addition to benefit to the public, researchers must describe benefit of the work to their own organisation and the data provider.**

**3.1.2. All projects should complete an acceptable form of ethical review – reviewing panels should include AI data ethics expertise where required.**

### 3.2. Safe People

**3.2.1. Research projects must have a named organisation taking responsibility for the design and conduct of the activity.**

**3.2.2. Researcher organisations:**

- Must sign an organisational agreement with the data provider on data use;
- Must be assessed for suitability of data access/use to demonstrate trustworthiness; and
- Should consider partnering with an academic or NHS institution.

**3.2.3. Researchers must:**

- Demonstrate relevant expertise and ability to support the project/data use;
- Complete supplementary training modules with specific content targeted at those using secure data without an academic research background
- Sign a user access agreement, countersigned by a senior member of their organisation.

**3.2.4. National organisations supporting TREs should consider maintaining a national unsafe user/organisation list on behalf of TREs.**

### 3.3. Safe Settings

**3.3.1. TREs should develop services to:**

- Increase approved software libraries;
- Allow virtual machines with varying configurations of hardware access and temporary internet access;
- Security risk assess new software and allow installation separately to data access;
- Support delegated software installation; and
- Isolate non-traditional researcher project spaces/virtual machines.

**3.3.2. TRE access should be restricted to organisationally managed machines and organisationally managed IP addresses only.**

## 3.4. Safe Data

- 3.4.1. Data providers should document any restrictions or caveats on use or users.
- 3.4.2. TRE staff should receive training in AI/ML techniques to understand data requirements and contributions to results, supplemented by expertise within review panels.
- 3.4.3. Disclosure risks (inbound and exbound) for novel/other types of data (e.g. qualitative, image, genomic) need to be identified and mitigated by TREs, data providers and researchers.

## 3.5. Safe Outputs

- 3.5.1. Risk assessment and mitigation of outputs from novel analysis techniques, tools, software and data need to be identified and understood by TRE staff and researchers. Safe output criteria should be developed between TREs for consistency.
- 3.5.2. Any TRE-specific disclosure control policies and requirements should be made clear to researchers on provision of data access.
- 3.5.3. A mechanism is needed to follow up on projects to determine how project results were used and benefits realised.

## 3.6. Overall

- 3.6.1. TRE governance documentation such as risk assessments should be amended to include specifics in relation to non-traditional researcher data access.

## 4. Project impact

---

### 4.1. Scotland impact

#### 4.1.1. Workstream 1 Governance Model

This model has been adopted by DataLoch acting as the Lothian Regional Safe Haven in Scotland. Its 5 Safe's structure allows for easy adoption in full or in part, by other Trusted Research Environments and will form a key part of the solution leveraged by the National Safe Haven in Scotland administrated by the eDRIS in the Public Health Scotland team.

The model has been shared across the Safe Haven network in Scotland, led by Research Data Scotland, and is being used as part of the wider RDS research to inform the future governance model for a federated Scotland solution.

#### 4.1.2. Workstream 2 Training Module Impact

The professional development training module has been designed to be consumable as a standalone service and is now available for integration into existing platforms for TRE users, current training providers, and data controllers across the DARE UK community. It is supplementary to existing accredited training, such as the MRC Safe Researcher training, and does not involve duplicated content. This means it can be consumed as an additional module for non-traditional researchers.

The learning platform will allow us to produce activity reports in various formats to track engagement by users and organisations e.g. number of user logins, activity and module completion rates, length of time in each sitting, evaluating the number of course enrolments, capturing user demographics and profiles etc. Through the analysis of learning system analytics, we can evaluate the learner performance and experience and improve the teaching content. A pre- and post-course survey will be embedded in the learning platform to understand user needs and refine and develop the teaching content based on user feedback. We are committed to creating a product that is

sustainable and has longevity and demonstrable impact. For this to be realised it is important to have insight into the benefits of the course for non-traditional users accessing TREs and identify barriers that may prevent them from engaging (or completing) the module. To this end we will embed a longer-term evaluation framework as part of this project to demonstrate impact.

### 4.1.3. Workstream 3 Public Engagement

Our engagement focused specifically on public perspectives in South-East Scotland around the security- and governance-access conditions required for non-traditional users to access sensitive data. The public consultation exercises have therefore directly helped to inform the design of DataLoch's information governance framework, particularly for non-traditional researchers. News articles on both the survey<sup>1</sup> and deliberative workshop<sup>2</sup> elements of this project (including access to the full reports) are available on the DataLoch website and add to the growing public feedback into safe and trustworthy data access.

## 5. Future Opportunities

---

- Further exploration of public perspectives that were common throughout the public engagement:
  - The public value and specific ethical considerations of AI and machine learning innovators;
  - Perspectives on differing trust levels for national and international research and innovation;
  - The perceived contradiction between profit and social value;
  - The influence of track records of both researchers and organisations in building or reducing perceived levels of trustworthiness.
- Pursue national accreditation and training alignment project, through engagement with relevant stakeholders and partners (e.g. SAIL, ONS, eDRIS) to align training delivery, standards, and achieve formal accreditation for the non-traditional researcher training module.
- Establish a national non-traditional researcher education working group and advisory service for the community of learners to share best practice and support.
- Establish a non-traditional researcher education community hub sharing regular information and updates on governance and data access requirements and a series of online webinars and networking events.

---

<sup>1</sup> <https://dataloch.org/insights/news/new-survey-results-public-perspectives-access-health-data>

<sup>2</sup> <https://dataloch.org/insights/news/defining-data-access-principles-local-residents>