

---

## Plan Overview

*A Data Management Plan created using DMPonline*

**Title:** UKRI Research Data Landscape Survey

**Creator:** Kieran Leach

**Principal Investigator:** Kieran Leach

**Affiliation:** University of Edinburgh

**Funder:** Engineering and Physical Sciences Research Council (EPSRC)

**Template:** EPSRC Data Management Plan

### **Project abstract:**

UKRI Research Data Landscape Survey

We are considering submitting a proposal to NetworkPlus intended to address the Community pillar and support delivery of the data roadmap deliverable. We hope to build on the methods and outputs of the Software Sustainability Institute's recent studies of research software to conduct a survey of UKRI funded research data which may be processed on future UKRI/DRI funded compute services.

In conducting such a survey we hope to gain visibility of the research data landscape, particularly in fields with which the traditional HPC community has not engaged. We would also hope to gain a better understanding of what data transfer technologies and tools are in use outside of the HPC community, and to communicate the data transfer experience the HPC community to those who may be able to make use of it.

**ID:** 177380

**Start date:** 01-06-2025

**End date:** 31-07-2026

**Last modified:** 09-05-2025

### **Copyright information:**

The above plan creator(s) have agreed that others may use as much of the text of this plan as they would like in their own plans, and customise it as necessary. You do not need to credit the creator(s) as the source of the language used, but using any of the plan's text does not imply that the creator(s) endorse, or have any relationship to, your project or proposal

# UKRI Research Data Landscape Survey

---

## Data Collection

### What data will you collect or create?

We will be collecting:

- Data required to conduct research studies (e.g. survey and interview data, data on research data held at institutions)
- Aggregated data sets summarising research data held at institutions

### How will the data be collected or created?

We will be collecting and generating data using:

- Forms, interviews and surveys
- Text and Data mining of existing data sources
- Analysis of existing data sources
- Documentation and Metadata

We will be adhering to the FAIR principles for data and software.

To maintain data quality, rigorous validation and testing procedures will be implemented during the data gathering and processing phases. Data reviews and audits will be conducted as required, and any anomalies or errors in the data will be corrected.

We will adhere to community and industry standards for data representation. This includes using version control systems, and employing metadata standards to enhance data interoperability.

## Documentation and Metadata

### What documentation and metadata will accompany the data?

The following outputs will contain all published data:

- A paper, or papers, containing collated information gathered by the survey;
- A report containing key findings and recommendations identified from the survey data;
- A poster for display at relevant workshops and submission to CIUK containing key findings from the survey and illustrating the diversity and geographical spread of the datasets identified.

Appropriate metadata will be included, and a digital object identifier will be included. These documents will be publicly available.

## **Ethics and Legal Compliance**

### **How will you manage any ethical issues?**

All data gathered will be of a type accepted by survey respondents as publicly publishable, and this will be validated as part of the survey process. Non-published data gathered during the process (such as PII of survey respondents) will be destroyed at the close of the project.

### **How will you manage copyright and Intellectual Property Rights (IPR) issues?**

All data gathered will be of a type accepted by survey respondents as publicly publishable, and this will be validated as part of the survey process. Respondents will be asked to further validate that there are no copyright concerns with their submissions being included in the published survey.

## **Storage and Backup**

### **How will the data be stored and backed up during the research?**

Data will be stored on University of Edinburgh provided storage services such as Microsoft Sharepoint or the University's Gitlab instance.

### **How will you manage access and security?**

Access will be limited to members of the project team via the controls available in University of Edinburgh provided storage services such as Microsoft Sharepoint or the University's Gitlab instance.

## **Selection and Preservation**

### **Which data are of long-term value and should be retained, shared, and/or preserved?**

Relevant data will be included in the outputs, and as part of publication remain available. EPCC will retain the original datasets used to create these outputs for at least 10 years, to support future potential work.

### **What is the long-term preservation plan for the dataset?**

Relevant data will be included in the outputs, and as part of publication remain available. EPCC will retain the original datasets used to create these outputs for at least 10 years, to support future potential work.

## Data Sharing

### How will you share the data?

Relevant data will be included in the outputs, and as part of publication remain available. EPCC will retain the original datasets used to create these outputs for at least 10 years, to support future potential work.

Relevant institutional, departmental and study policies on data sharing and security include:

Policy	URL or Reference
Data Management Policy & Procedures	<a href="https://information-services.ed.ac.uk/about/policies-and-regulations/research-data-policy">https://information-services.ed.ac.uk/about/policies-and-regulations/research-data-policy</a>
Data Protection Policy	<a href="https://data-protection.ed.ac.uk/data-protection-policy">https://data-protection.ed.ac.uk/data-protection-policy</a>
Data Security Policy	<a href="https://infosec.ed.ac.uk/information-protection-policies/information-security-recommended-reading">https://infosec.ed.ac.uk/information-protection-policies/information-security-recommended-reading</a>
Data Sharing Policy	<a href="https://information-services.ed.ac.uk/about/policies-and-regulations/research-data-policy">https://information-services.ed.ac.uk/about/policies-and-regulations/research-data-policy</a>
Institutional Information Policy	<a href="https://information-compliance.ed.ac.uk/freedom-information/published-information/information-compliance">https://information-compliance.ed.ac.uk/freedom-information/published-information/information-compliance</a>
DataShare service policies	<a href="https://library.ed.ac.uk/research-support/research-data-service/after/data-repository/service-policies">https://library.ed.ac.uk/research-support/research-data-service/after/data-repository/service-policies</a>

### Are any restrictions on data sharing required?

All data collected will be published, with the exception of non-published data gathered during the process (such as PII of survey respondents) which will be destroyed at the close of the project.

## Responsibilities and Resources

### Who will be responsible for data management?

The Data Manager for this activity will be confirmed at project start.

### What resources will you require to deliver your plan?

No resources are required to deploy this data management plan.