

# Guiding principles for sharing data in the University of Leeds Library Data Repository

The Research Data Leeds (RDL) Repository is the institutional research data repository for the University of Leeds. The service aims to facilitate open research, data discovery and data sharing. The Restricted Access Data Repository (RADAR) is a companion service to the Research Data Leeds repository. RADAR provides restricted access (via a download model) to data that cannot be made fully open.

## What is research data?

Research data refers to any information collected, observed, generated, or created to support and validate original research findings. While often digital, it can also include non-digital formats, such as laboratory notebooks or diaries.

Importantly, sharing data does not always mean making it fully open. Certain data may need to remain restricted due to confidentiality, sensitivity, or legal agreements, such as non-disclosure agreements.

## Human-derived data and anonymisation

For human derived data, anonymisation is typically required before it can be deposited. Anonymised data is no longer considered personal data and is not covered by the Data Protection Act or UK GDPR ([ICO](#)).

Once data is anonymised, it can be shared without requiring participants' permission. However, for ethical reasons, and whenever possible, it is important to inform participants about any plans to share their anonymised data ([UK Data Service](#)). The participant information sheet should provide details about any plans for sharing the anonymised data ([UK Data Service](#)).

Templates for participant information sheets and consent forms can be found on the [University of Leeds Research Ethics webpages](#).



# Guiding principles for data deposit

The Library Research Data Service has developed a set of principles to guide decision-making around data deposit requests. These principles, which have been endorsed by the Data Access and Retention Group (DARG), aim to ensure that deposits align with ethical, legal, and institutional requirements.

If agreement cannot be reached between the Library Research Data Service and the researcher on any of these principles, the decision on whether to accept the deposit or what level of classification the dataset will be given will be escalated to the Data Access and Retention Group for a decision.

DARG has representatives from Information Governance, Library Research Services, Data Services, Research Ethics, Leeds Institute of Health Sciences and Contracts (Research and Innovation Services).

## 1.Data should be effectively anonymised and the risk of reidentification should be low

It is the researcher's responsibility to ensure that anonymisation has been carried out thoroughly, to the extent that risk of reidentification has been reduced sufficiently to a remote level and the data is 'effectively anonymised' ([as outlined by the ICO](#)).

Library Research Data colleagues will review the data to check for any apparent issues, eg any obvious names, addresses etc. However, since they are not anonymisation experts, they cannot provide any guarantees regarding the effectiveness of the anonymisation. If participants have been told that their data will be anonymised but the researcher cannot provide these assurances, then the data cannot be deposited in the University of Leeds Data Repository.

When researchers deposit data in the Library Research Data Repository, they must sign the Data Deposit Agreement, confirming whether participants were informed their data would be anonymised and that they have ensured effective anonymisation. If the data is pseudonymised, the researcher will need to provide assurances that the pseudonymisation has been carried out appropriately and the key is secured or has been destroyed.

## 2.If data cannot be anonymised, eg videos of performances involving people, or if the researcher intends to share identifiable data, participants must have given permission for the data to be shared either on an open or restricted basis

If a researcher wants to deposit identifiable data in the University of Leeds Data Repository, it is their responsibility to provide evidence that participants have given permission for their non-anonymised data to be shared on a restricted or open basis. This would usually be in the form of participant information sheet wording and consent forms.

A privacy statement should also have been provided to participants, outlining that the legal basis for data processing is "public interest."

Links to the University Privacy statement is included in the [University Participant Information Sheet Template](#).

### 3.If the dataset contains special category data or other sensitive information such as commercial data, data sourced from a third party or data that could pose a risk to individuals, extra caution is needed

[Special category data](#), as defined in the UK GDPR, requires additional care and caution. It is the researcher's responsibility to inform participants how their special category data will be used, and this should be evidenced in the participant information sheet wording and consent forms.

If the data contains any special category or sensitive data, the Library Research Data Service will discuss with the researcher whether any of the data could be removed from the dataset and if not, whether it is appropriate to share this data on an open or restricted basis.

If data has been generated with or sourced from third party project partners, including commercial partners, the Library Research Data Service will ask to see any necessary data sharing agreements, to ensure there are no restrictions on sharing the data. For example, this could include data sourced from existing primary research or analysis of human biological materials sourced from a tissue bank. More information on safeguarding data, including things to consider before sharing or reusing data can be found on the [Library's Safeguarding webpages](#).

Data that has been scraped from social media and other online sources and which is not anonymised will also require additional consideration to assess the reasonableness of sharing it on an open basis.

### 4.If the data does not contain any special category or other sensitive data and it is effectively anonymised, even if participants haven't been fully informed that their data will be shared openly, the data would usually be accepted into the repository and shared openly

Such data would usually be shared openly, as long as the data is effectively anonymised and ethical permission to share the data has been sought. The Library Research Data Service will discuss with the researcher.

#### **If the participants have been told explicitly that their anonymised data won't be shared, then extra caution is needed**

If the participants have been told explicitly in the participant information sheet that the anonymised data collected from them would not be shared, either openly or on a restricted basis, the Library Research Data Service will discuss with the researcher whether it is appropriate for the data to be shared. Wording could include 'data will be destroyed' or 'data will be kept strictly confidential'.

### 5.If the participants have been given inconsistent or contradictory information about what will happen to their data, or they have been told their data will only be shared for future research, extra caution is needed

If the participants have been given contradictory or inconsistent information in the consent forms and participant information sheet, the Library Research Data Service will review the information and discuss with the researcher whether it is appropriate to share the data and escalate to DARG if necessary.