

Prompt to Good Data Sharing

Much data - from or about people - can be shared ethically and legally post-project. Exactly what data can be shared and how requires thought from the beginning of your project.

Don't assume that if all of your data cannot be fully open then nothing can be shared – make plans so that your data can be shared with access “as closed as necessary and as open as possible”.

There is an expectation that post-project data is stored centrally, even for confidential data which can only be accessed by the research team.

Highly sensitive but non-linked pseudonymised material can be made available “on request” to other bona fide researchers using non-disclosure data sharing agreements. Please consider the following when planning your [Data Management](#):

- ethics
- participant information sheets
- consent forms
- storage
- controlling access

Before data collection

You will need ethical approval if you are collecting data from humans or analysing data relating to individuals. Please consider the points below when preparing your [consent](#) and [participant information](#) forms. Although you should not ask for consent to share data as that is not the University's legal basis for undertaking research, it is good ethical practice to let participants know what will happen to the research data after the project is finished.

Please use the Consent form to seek their understanding for how the data will be stored and used for future research and learning. (Please see [ERGO](#) templates for Consent and Participation Information forms)

Please use the participant information form to outline the steps taken to ensure data can be shared during and after the project.

- Explain the benefits of data sharing indicating whether research data will be deposited in a data repository, naming the organisation responsible for holding the data. · Specify whether the data will be anonymised and how.
- In what form will the data be held e.g. anonymised transcripts, audio recordings, survey database. Whatever the form the data is in, it should be unlinked from any pseudonymisation keys.
- Specify where all the data (including contact details) will be held during the project and where the non-linked data will be deposited after the project.
- Specify whether use or access restrictions will apply to the data in the future e.g. exclude commercial use, safeguarded access.
- State clearly that de-identified data can be used in research publications. Remember that participants need to know their anonymised data may be shared with other researchers as well as the results published – you can't ask them afterwards!
- An example statement you may use in the consent form is as follows:

I understand that [specify the personal information and/or special category data to be stored] that I provide may be deposited to [name of repository]/ held by (name of researcher/organisation) as described in the participant information sheet so it can be used for future research and learning

Look at the [Data Ethics Canvas](#) from the Open Data Institute, a useful tool to help you identify and manage ethical issues.

De-identifying data

You can anonymise OR pseudonymise your data. Both allow sharing. Please see the library guide on [Sensitive Data and GDPR](#).

Anonymised data does not fall under the scope of GDPR so easier to use but can be difficult to remove all identifiers without losing richness from the data. This is where data has been permanently de-identified.

Pseudonymised data still falls under GDPR but is much more likely to be attainable. This is where direct identifiers (name, age, place etc) have been removed and replaced with codes. Context is key when dealing with de-identifying data. A dataset can be functionally anonymous if transferred from NHS without research key but could “in theory” be re-identified by linking to other sources. So, assume all data is pseudonymous. Please see the [ICO Anonymisation Code of Practice](#) for more information; p22 Motivated Intruder Test.

We also recommend this very useful guide on [how to anonymise qualitative and quantitative data](#) from the UK Data Service.

Storing your data during the project

- All University research should be stored on University servers.
- Keep your research data in your networked My Documents, the Research Filestore or in your University account OneDrive or Sharepoint
- Do not use non-University cloud services (for example Dropbox, Google Drive) to store research data. Do not email non-anonymised research data or store it on usb sticks.
- If you have sensitive personal data (see below for more information) then you should be extra careful and may want to password protect and encrypt your files as well. Use SafeSend and not email if you need to send another researcher your data.
- Do not send data to anyone before it has been pseudonymised or anonymised. An exception would be if there is a legal contract and appropriate protection/safeguards in place, for example for transcription or data entry by a third party.
- Do not store the anonymisation/pseudonymisation key in the same place as the data.
- Do not share the anonymisation/pseudonymisation key with anyone.

Ready to Share your Data?

Controlling access - You can safeguard access to your confidential and sensitive data in a 2-step process.

First confirm conditions under which the data can be released:

- Bona fide researchers with ethical approval may request data from the University of Southampton Repository
- Bona fide researchers with ethical approval who have agreed to the usage/data access agreement will be given access
- Anyone who has agreed to the usage/data access agreement will be given access [Note the University does not currently have a generic data sharing agreement. The research team who holds the data must decide on conditions of access]
- A custom condition – please supply

Second, confirm who can make the decision:

- The Library Research Data Service
- Specific people (please specify) - typically one or more of the dataset creator(s), and / or their supervisor (for PGRs)
- If all of these people have left the university the request will be escalated within the faculty as appropriate.
- Another arrangement – please specify.

Normally the requestor would have to demonstrate they have ethics clearance from their home institution (which the Library then double checks with the institution) and have signed a data sharing agreement stating conditions for how they use the data, for example not republishing or further sharing the dataset, destroying the dataset after their project and not seeking to re-identify the subjects. This is becoming standard practice across most universities in the UK.

[Please see these examples](#)

University of Southampton

Dataset for: An exploration of factors influencing parents from minoritized ethnic backgrounds views and experiences of accessing perinatal mental health services in the UK Life

<https://doi.org/10.5258/SOTON/D3149>

On request for ethical reasons, library can make decision.

Dataset for: Digital health technologies in improving efficiency in reproductive medicine

<https://doi.org/10.5258/SOTON/D2797>

One file on request for ethical reasons, the rest are open

Dataset for: Developing SERS-based point-of-care inflammatory monitoring for risk identification of worsening progression of age-related disease

<https://doi.org/10.5258/SOTON/D2222>

On request for ethical reasons supervisor to approve requests

USEWOD 2016 Research Dataset

<https://doi.org/10.5258/SOTON/38534>

On request as a usage agreement needs to be filled out.

Externally

Openly shared de-identified data such as the [Pregnancy and Lifestyle Study](#) in Australia

Mediated access (ask for permission) Dataset for the [Epidemiology of Chronic Ankle Problems in the community](#) from the University of Sydney.

[Storing Data after the project](#)

The University of Southampton Research Data Management Policy has a requirement that all **significant** Research Data should be held for a minimum of 10 years and may be longer where the data is actively used. Funders also have retention requirements and some research data will also be subject to legal requirements. This does not apply to UG or PGT students. (please see addendum for Health Sciences below)

If you are unsure about what to deposit and where, see the Library guidance on what is [significant data](#) and where to [deposit](#).

What are the benefits to me?

- can increase your citations
- allows data to be tested and validated
- allows data to be re-used for further research or in teaching

increasingly institutions/funders see sharing data as in the public interest in line with the [OECD](#) Principles and guidelines for Access to Research Data from Public Funding

- [UKRI data policy](#)
- [Wellcome Data sharing](#)
- reduces duplication of effort and costs
- Researchers in developing countries can see your work.

The National Institutes of Health's [Policy for Data Management and Sharing \(DMS Policy\)](#), went into effect 2023. The policy includes an expectation that researchers will maximize their data sharing within ethical, legal, or technical constraints, and explicitly encourages researchers to incorporate data sharing via deposit into a public repository into their standard research process.

Addendum

21/11/2019: School of Health Sciences Ethics Chair states that *"To mitigate against the situation where data are published, but not kept, we typically advise all students (including UG and PGT) to state that their data will be kept for a minimum of 10 years."*

Links

Help from the Research Data team

- <http://library.soton.ac.uk/researchdata/planning>
- <http://library.soton.ac.uk/researchdata/storage>
- <https://library.soton.ac.uk/researchdata/sharing>
- <http://library.soton.ac.uk/researchdata/gdpr>
- Email researchdata@soton.ac.uk

Patients/Sensitive Data

- [Publishing sensitive data flow chart](#)

Policies

- <https://www.ukri.org/manage-your-award/publishing-your-research-findings/making-your-research-data-open> - UKRI- UKRI
- <https://www.southampton.ac.uk/~assets/doc/calendar/Research%20Data%20Management%20Policy.pdf> - UoS
- <https://wellcome.org/grant-funding/guidance/policies-grant-conditions/data-software-materials-management-and-sharing-policy> - Wellcome Trust

Sharing is good!

- <https://vimeo.com/125783029> - Open Data Institute
- <https://www.nature.com/articles/d41586-019-01506-x> Career prospects
- <https://safesend.soton.ac.uk/>