A Guide to the Research Data Service
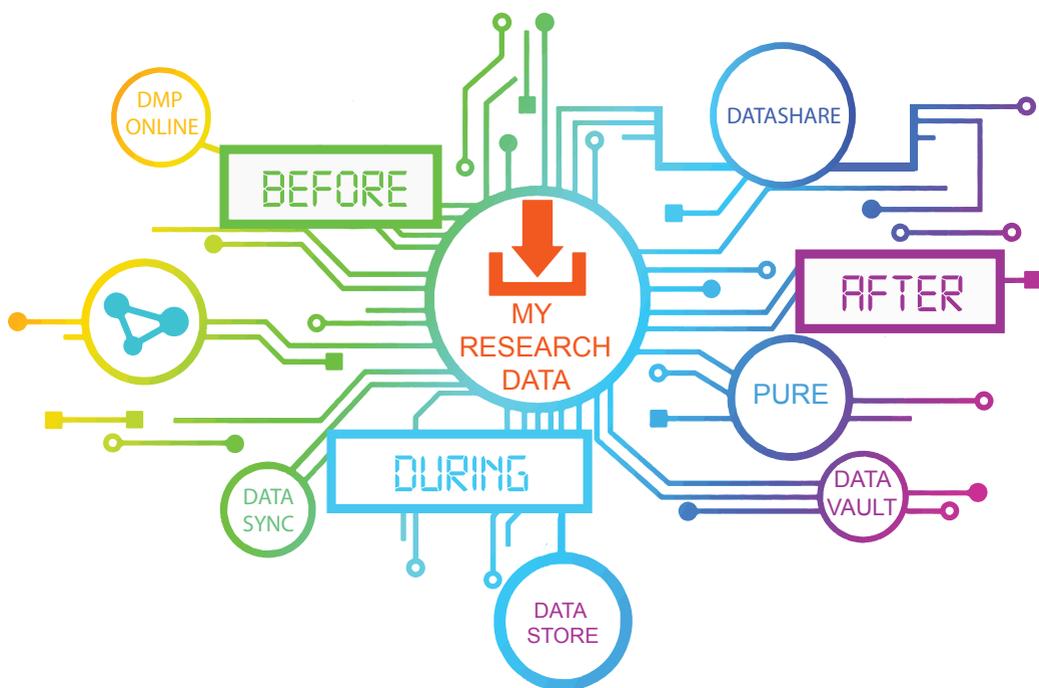
THE UNIVERSITY *of* EDINBURGH

Table of Contents

This booklet was produced in August 2018 by the Research Data Support Team, Library and University Collections, Information Services, University of Edinburgh, data-support@ed.ac.uk.

The latest version can be downloaded from:

https://www.ed.ac.uk/is/research-support/research-data-service/about-the-research-data-service

Research Data Management

Research Data Management (RDM) is a general term covering how you organise, structure, store, and care for the data used or generated during the lifetime and after the completion of a research project.

In addition to many major funders requiring research data to be actively managed, It is good research practice to ensure that your data are managed properly throughout the life of the project. This means planning how you will collect, store, and care for your data before you start the research process, through to how you will ensure it is maintained in the longer-term, and, if appropriate, shared with your research community and beyond.

The University of Edinburgh has a formal Research Data Management Policy that requires research data to be managed to the highest standards as part of the University's commitment to research excellence. The ensures that sensitive data are suitably protected, and the non-sensitive data reach the widest possible audience.

Information Services provides a suite of tools and support to assist you in managing your data according to policy and best practice. This guide provides an overview of the tools, explaining what they are, why you should use them, and how to use them.

University of Edinburgh Research Data Management Policy:
https://www.ed.ac.uk/is/research-data-policy
Research Data Service:
http://www.ed.ac.uk/is/research-data-service

Research Data Service

We provide and advise upon a range of tools and support to help you before, during and after your research project:

BEFORE: identify existing datasets for reuse, plan how you will collect, manage, and store your data, including storage and backup, sensitivity of the data, and whether it can be shared;

DURING: analyse data, store and backup active data, synchronise your data across devices and with your research collaborators, and version your software and files;

AFTER: record, share, and archive your data for the long-term;

TRAINING AND SUPPORT: online and face-to-face workshops and courses, plus direct support when you need it.
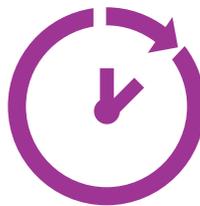
BEFORE	DURING	AFTER	TRAINING & SUPPORT
create a data management plan	working with data	share and archive your data	

Creating a Data Management Plan

What is DMPonline?

DMPonline is a tool created by the UK's Digital Curation Centre (DCC), based here at the University, which helps researchers write and maintain data Management Plans (DMPs).

Why should I use it?

The tool includes a number of templates for funders in the UK and overseas to help write your DMP according to the specific requirements you need to meet. It is customised for Edinburgh use too so you are provided with examples and guidance based on local support and services. If you are working with other researchers you can share your plan to write collaboratively. You can also export your plan once you complete it. You can view your plan as a web page and also download it as a plain text or as a PDF. Upon completion the plan can be downloaded in a variety of formats to enable submission alongside a grant application.

How can I access it?

The tool is available from: http://dmponline.dcc.ac.uk/. You can sign up by entering your email, organisation, and password. You can also login using your EASE username and password once you have linked these institutional credentials to your DMPonline account.

Contact the Research Data Support Team at data-support@ed.ac.uk to assist you with your data management plan.

BEFORE

Finding Data

What is the Data Library Portal?

The Research Data Support Team assists researchers to discover and use datasets for analysis, learning and teaching.

Why should I use it?

Your research project may involve acquiring and analysing existing datasets, not just collecting new data. In some cases you may know where to find these datasets; in others you may need help to find or to access them. Some data may have special access requirements for reasons of confidentiality, or be behind a paywall.

The Data Library can help you find data freely available on the internet or licensed by the University. Data resources include surveys, censuses, mapping, and databases.

We can help you create teaching datasets and provide hands-on data workshops for your classes.

How can I access it?

View the Finding data page:
https://www.ed.ac.uk/is/finding-data

DURING

Storing Data

What is DataStore?

DataStore provides file storage for active research data, and is available to all research staff and postgraduate research students (PGRs).

DataStore provides a free individual allocation for each researcher, as well as shared group spaces. For more demanding projects additional capacity of virtually any size can be arranged.

Why should I use it?

It is fully backed-up, secure, resilient, multi-site storage, and new storage allocations can be created immediately.

How can I access it?

You can connect to DataStore as a network shared drive, and users of the University's Supported Desktop already have their individual DataStore space mapped to their M: drive.

Other connection methods and full connection instructions are available at: http://www.ed.ac.uk/information-services/computing/desktop-personal/network-shares

DURING

Sensitive Data

What is Data Safe Haven?

Data Safe Haven is a controlled and secured service environment for undertaking research using sensitive data (personal, sensitive-personal, or confidential). The service provides robust controls and safeguards to enable the secure transfer of sensitive data into a 'walled garden' environment where it can be securely stored, manipulated and analysed by approved members of a research team.

Why should I use it?

- To enable you to access, and work with sensitive data that requires a high degree of security.
- To allow a team of authorised researchers to work together on a project utilising sensitive data.
- To meet stringent requirements of data owners for access to their data (e.g. National Health Service).

How can I access it?

If you think your research may require the use of a Data Safe Haven then you should contact the Research Data Support team before submitting a grant application, as there may be costs involved in setting up a suitable environment for your project which must be included in the project budget.

Email us at: data-support@ed.ac.uk

DURING

Collaborating

What is DataSync?

Datasync is a tool to allow researchers to share data with collaborators and to synchronise data across multiple devices. The web interface can be used to share data with anyone who has a valid email address and allows access from any web browser. It also has an app to synchronise data between computers and mobile devices.

Why should I use it?

You will have dedicated DataSync storage, plus the ability to connect up your personal and group data on the RDM DataStore for effectively unlimited storage capacity, giving you the ability to synchronise and share your active research data.

How can I access it?

You can activate your account by logging in to the DataSync web site using your University username and password at: https://datasync.ed.ac.uk

Apps are available for Windows, Mac OS X, Linux, iOS and Android.
Documentation is available at: http://www.ed.ac.uk/is/datasync

DURING

Version Control

What are Subversion & GitLab?

Subversion & GitLab are version control tools which allow users to store and track changes to code and documents. Subversion is also available as an extension called SourcEd which provides a web-based collaboration tool integrated with your repository. GitLab is a web frontend to Git that provides additional features, including issue management.

Why should I use it?

When files stored in a Subversion or Git repository are updated, older versions are retained, making it possible to revert to them if required. The service also allows for multiple people to collaborate on a project, tracking multiple changes to source code and documents authored by multiple contributors.

How can I access it?

Subversion documentation is available at:

http://www.wiki.ed.ac.uk/display/ResearchServices/Version%2BControl%2B

Gitlab documentation is available at:

http://www.wiki.ed.ac.uk/display/ResearchServices/Version%2BControl%2B

The GitLab service can be accessed at:

https://git.ecdf.ed.ac.uk/

DURING

Recording Datasets

What is PURE?

PURE is the current research information system. It contains individual profiles of research information for all staff members involved in research. Staff members are encouraged to add content and edit their profiles with details of projects, publications, datasets, etc.

Why should I use it?

The University is using PURE to record descriptive data (metadata) about research data in order to meet institutional and funder policy requirements. The metadata must be sufficient to allow others to understand what research data exists, how it was generated, and how to access it. Datasets catalogued in Pure are automatically included in a staff member's online profile in Edinburgh Research Explorer (http://www.research.ed.ac.uk) alongside other research outputs such as journal articles, conference papers, and research grants.

Whether data are stored in DataShare or an external repository, a persistent identifier, such as a DOI, can be added to the Pure record to link them to grants and publications creating a more robust scholarly record.

How can I access it?

Users can log into PURE using their EASE Username and password via:
http://www.pure.ed.ac.uk/
For information and help on how to add a dataset description to PURE see:
https://www.ed.ac.uk/information-services/research-support/research-information-management/pure

AFTER

Sharing Open Data

What is DataShare?

Edinburgh DataShare is a free, open data repository which allows University of Edinburgh researchers to upload, share, and license their data resources for online discovery and re-use by others. Your data will be discoverable through search engines to maximise visibility and impact. DataShare can provide you with usage statistics so you know when your data have been downloaded.

Why should I use it?

- Increase impact of your research;
- Make your data assets discoverable and citable;
- Minimise the risk of data loss for your own future use;
- Meet funding body requirements & comply with University policy;
- Ensure your research data is preserved for posterity.

A big advantage of depositing your data is that they will be preserved - even for your own future use. The system creates a permanent record, a persistent identifier (DOI), and a suggested citation, so that your work can be formally attributed when re-analysed by others.

How can I access it?

Simply click on the Deposit button on the home page of the repository and log in with EASE: http://datashare.is.ed.ac.uk

AFTER

Archiving Data

What is DataVault?

The DataVault is an archive storage service where you can safely store your research data for the long term.

Why should I use it?

Your data will be kept safe from accidental deletion or inappropriate access and, when combined with a record of the dataset in PURE, will fulfil the expectations of research funders who require long term storage of research data.

Data can be copied into the DataVault from your DataStore, and when a retrieval request is made, a copy will be placed back into your DataStore. Using the DataVault will allow you to archive old data from your DataStore, to freeing up space for new research. The DataVault also makes a good location for storing specific versions of your data, for example at the end of a grant.

How can I access it?

DataVault is a chargeable, web-based system, accessible using your EASE username and password. User guides and other information can be found at: https://www.ed.ac.uk/information-services/research-support/research-data-service/sharing-preserving-data/datavault

AFTER

Training - Online

MANTRA

MANTRA is an open, web-based training course intended for self-paced learning by PhD students, early career researchers, or anyone else who manages digital research data. It provides guidance about good RDM practice, illustrated with real-world stories.

You can take the course from start to finish, or pick and choose from nine data management topics, including data management planning, data protection, documentation and metadata, and storage and security.

MANTRA website: https://mantra.edina.ac.uk/

RDMS MOOC

The Research Data Management and Sharing MOOC (Massive Open Online Course) uses the Coursera on-demand format to provide short, video-based lessons and assessments. It is normally taken across a 5-week period with a cohort of other learners. Certificates are available to any learner who completes a course for a small fee. Subjects covered are:

- Understanding Research Data
- Data Management Planning
- Working with Data
- Sharing Data
- Archiving Data

The MOOC has been developed and delivered by the University of Edinburgh in collaboration with the University of North Carolina:

http://www.coursera.org/learn/data-management

TRAINING
& SUPPORT

Face to Face

Workshops and courses

We provide short courses and interactive workshops open to all research staff and postgraduate research students interested in finding out how to manage their research data effectively and efficiently. They focus on issues such as good practice in research data management, working with personal and sensitive data, writing data management plans, and handling data with SPSS.

Courses on any aspect of RDM can be tailored for schools, institutes or research groups on demand. Here is a selection of topics that we currently offer:

- Managing your research data: why is it important and what should you do?
- Working with personal & sensitive research data
- Creating a data management plan for your grant application
- Good practice in research data management
- Handling data using SPSS

Awareness raising

Bespoke RDM sessions for Schools/Centres/Institutes are organised and delivered by the Research Data Support Team.

https://www.ed.ac.uk/information-services/research-support/research-data-service/training

TRAINING & SUPPORT

Support

Research Data Service Website

The Research Data Service website is a one stop shop for information about research data management and related services available at the University of Edinburgh:

https://www.ed.ac.uk/is/research-data-service

Edinburgh Research Data Blog

The Research Data Service blog has been set up to communicate and inform about RDM activities and developments at the University:

http://datablog.is.ed.ac.uk

Help and Support

All enquiries about research data management can be made to IS Helpline or directly to Research Data Support:

data-support@ed.ac.uk
Telephone: +44 (0)131 651 5277

If you require this document in an alternative format, such as large print or a coloured background, please contact The IS Helpline by email IS.Helpline@ed.ac.uk or by telephone +44 (0)131 651 5151