



# Research Data Management

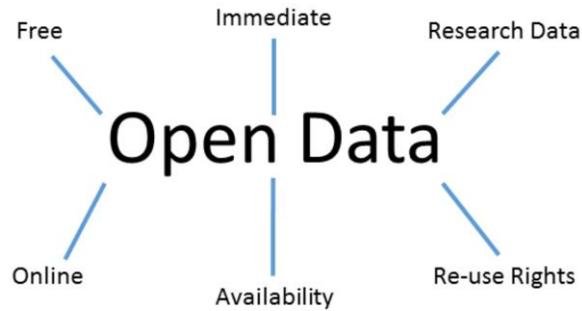
Dr. Chris Tibbs  
Research Data Officer  
[rdm@exeter.ac.uk](mailto:rdm@exeter.ac.uk)



# Open Data

---

- **Open Data** = a new paradigm across all disciplines, providing benefits to individual researchers, institutions, and funders
- Good research data management habits are essential to creating data that are suitable for sharing



Open Data is a new paradigm in which research data are freely and openly shared, with full re-use rights. Open data helps to maintain research integrity and enables validation of results. Additionally, open data allows the public to access publicly funded research.

The main point to remember about the open data environment is that data preservation is key, however, to ensure that data are suitable for sharing, good research data management practices are essential.

# What is Research Data Management?

---

- **Research data** = all information collected/created and used in research
  - structured data (e.g., databases, tables, etc.)
  - unstructured data (e.g., textual sources, images, audio recordings, personal notes, emails, etc.)
- **Data management** = effective management of the information lifecycle
  - planning for data
  - working with data
  - preserving data

Research data includes *\*all\** of the information that is used for research, irrespective of its format.

Research data management can be broken down into 3 key aspects:

- good research data management requires advanced planning prior to the start of a project
- you need to implement effective day-to-day data management practices during a project
- you need a long-term strategy for data preservation after the project is completed

## **Importance of Good Data Management**

---

- Data are the cornerstone of research
- Good quality data leads to good quality research
- To protect data from loss, destruction, corruption
- Ensure that data remain accurate and reliable
- Increase research productively
- Enables compliance with ethical codes, data protection laws, journal requirements and funder/institutional policies



## UoE Requirements – Key Points

---

- Research data management is the joint responsibility between PI(s) and researcher(s)
- All research proposals must include data management plans
- Research data management is a valid cost for research proposals
- Researchers are responsible for ensuring that data are deposited in an appropriate repository after project completion
- Publications should include a short data access statement

Full details: <http://hdl.handle.net/10871/26168>



## Funder Requirements – Key Points

---

Funders are increasingly requiring researchers to meet certain data management criteria

- Submission of a technical or data management plan when applying for funding
- Open data sharing after project completion
  - Deposit data in a data repository
  - Minimal or no access restrictions
- Long term preservation of the data
  - Most funders require 10+ years

Further details on specific funder policies: <http://v2.sherpa.ac.uk/juliet/>



# Data-related policies



● Full Coverage ● Partial Coverage ○ No Coverage

Research Funders	Policy Coverage			Policy Stipulations				Support Provided			
	Published outputs	Data	Time limits	Data plan	Sharing/access	Long-term curation	Monitoring	Guidance	Repository	Data centre	Costs
AHRC	●	●	●	●	●	●	○	●	○	●	●
BBSRC	●	●	●	●	●	●	●	●	●	●	●
EPSRC	●	●	●	●	●	●	●	●	○	○	●
ESRC	●	●	●	●	●	●	●	●	●	●	●
MRC	●	●	●	●	●	●	○	●	●	○	●
NERC	●	●	●	●	●	●	●	●	●	●	●
STFC	●	●	●	●	●	●	●	●	●	●	●
Cancer Research	●	●	●	●	●	●	●	●	●	○	●
European Commission	●	●	○	●	○	○	○	●	●	●	●
Wellcome Trust	●	●	●	●	●	●	●	●	●	●	●

<http://www.dcc.ac.uk/resources/policy-and-legal/overview-funders-data-policies>

This table displays the major research funders' policy coverage for a range of different aspects of research data management.

## UKRI Common Principles

UK Research  
and Innovation

- Publicly funded research data are a public good, produced in the public interest, which should be made openly available with as few restrictions as possible in a timely and responsible manner.
- To enable research data to be discoverable and effectively re-used by others, sufficient metadata should be recorded and made openly available to enable other researchers to understand the research and re-use potential of the data. Published results should always include information on how to access the supporting data.
- It is appropriate to use public funds to support the management and sharing of publicly-funded research data. To maximise the research benefit which can be gained from limited budgets, the mechanisms for these activities should be both efficient and cost-effective in the use of public funds.

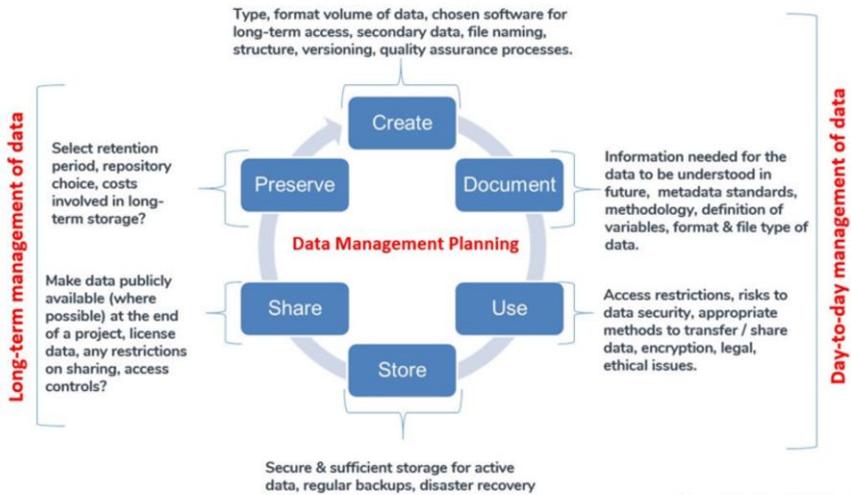
<http://www.rcuk.ac.uk/research/datapolicy/>

UKRI have a set of principles to guide the data policies of the 7 UK Research Councils.

# Data Management Planning



# Research Data Lifecycle



© Stuart Macdonald/EDINA. Used with permission



## **What is a data management plan (DMP)?**

---

A DMP is a formal statement describing how research data will be managed and documented throughout a research project and the terms regarding the subsequent deposit of the data in a data repository for long-term management and preservation.

- What data will be collected/created (format, types, and size) and how?
- How will the data be documented and described?
- How will data ethics and Intellectual Property be managed?
- What are the plans for data sharing and access?
- What is the strategy for long-term data preservation?

A data management plan (DMP) is a basic statement describing how the research data will be managed throughout a project and beyond.

A data management plan is a valuable tool for navigating a research project, but should not be viewed as set in stone. In fact, it's good practice to review your plan periodically, and revise or update it as necessary.

# DMP Online

DMP Online is a web-based tool, developed by the Digital Curation Centre (DCC), to help write data management plans

<https://dmponline.dcc.ac.uk>

A screenshot of the 'Create a new plan' page on the DMP Online website. The page has a grey header with the 'DMP ONLINE' logo and navigation links: 'View plans', 'Create plan', 'About', 'Future plans', 'Help', and 'Change language'. The main content area is white and contains the following sections:

- Create a new plan**: A heading followed by a sub-heading 'What research project are you planning?' and a text input field for 'Project title'. A note below the field says 'if applying for funding, state the title exactly as in the proposal.'
- Primary research organisation**: A section with the heading 'Select the primary research organisation responsible'. A dropdown menu shows 'University of Exeter' with a search icon. Below it is a checkbox labeled 'My research organisation is not on the list or no research organisation is associated with this plan'.
- Funding organisation**: A section with the heading 'Select the funding organisation'. A text input field is followed by the text 'Begin typing to see a filtered list'. Below it is a checkbox labeled 'No funder associated with this plan'.

A 'Create Plan' button is located at the bottom of the form.

DMP Online is a web-based tool that contains data management plan templates for all of the major research funders, and provides guidance and suggestions for what to include. It also contains a data management plan checklist and a range of example data management plans.

## What does DMP Online do?

---

A web-based tool that enables users to

- i. **Create, store, and update** multiple versions of DMPs throughout the research lifecycle
- ii. **Collaboratively work on the DMP** with multiple people able to access the DMP
- iii. **Meet a variety of specific data-related requirements** (from funders, institutions, etc.)
- iv. **Get tailored guidance** on best practice and helpful contacts, at the point of need
- v. **Customise, export, and share DMPs** in a variety of formats in order to facilitate communications within and beyond research projects



# Preserving And Sharing Data



## FAIR Data

---

To meet the funder requirements, you should make your data:

**F**indable   **A**ccessible   **I**nteroperable   **R**eusable



Image: CC BY-SA 4.0 Sangya Pundir

<https://www.force11.org/group/fairgroup/fairprinciples>

In general, your research data should be FAIR. Findable means that the data should be registered in a searchable repository and be assigned a persistent identifier. Accessible means that the data must be available via a standardised protocol e.g., http, and that the metadata is available, even if the data themselves are not. Interoperable means that the data must be able to be seamlessly integrated with existing data and tools, and Reusable means that the data must be clearly licensed.

## Research Data Repositories

---

- Digital research data are best preserved and published using a research data repository
- A repository is an online database service, an archive that manages the long-term storage and preservation of digital resources and provides a catalogue for discovery and access
- Most data repositories do not charge to deposit research data, though many require registration

Registry of research data repositories: <http://www.re3data.org/>

Although most funders require data to be preserved in a repository, not all of them provide a specific repository, simply allowing the data to be deposited in any appropriate repository.

Subject based repositories are available, and can be found using the registry of research data repositories.

# Open Research Exeter (ORE)

ORE is the University of Exeter's institutional repository, suitable for both publications and data

ORE offers long-term data preservation and curation

Runs on DSpace software and is currently undergoing an upgrade from DSpace 4.2 to 5.7

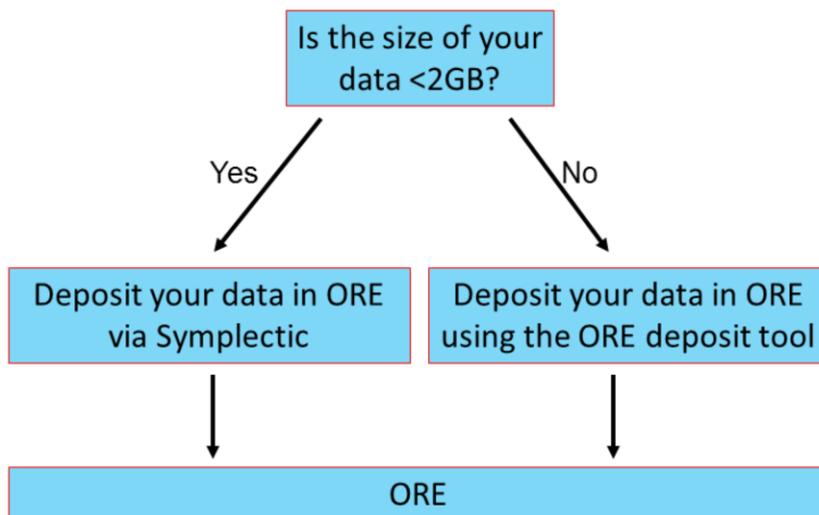
<https://ore.exeter.ac.uk/repository/>

The screenshot shows the homepage of the Open Research Exeter (ORE) website. At the top, there is a navigation bar with the University of Exeter logo and a search bar. Below the navigation bar, the main content area features a search bar, buttons for 'Upload to ORE' and 'Login to ORE', and a 'Recently Uploaded' section. The 'Recently Uploaded' section lists several research articles with their titles and authors. On the right side, there are 'Related Links' and 'Browse' sections.

ORE is a showcase of the research outputs of the University.

## Depositing Data In ORE

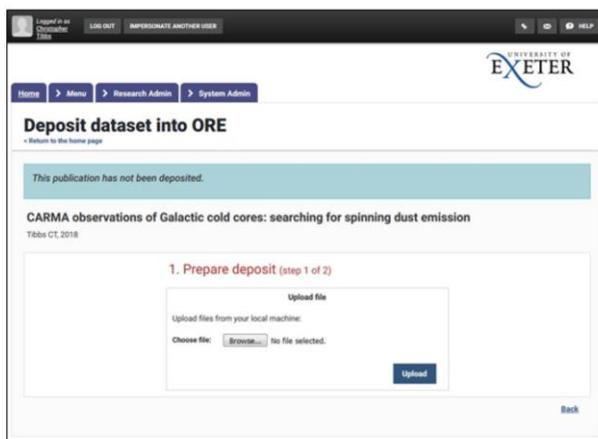
---



## Depositing Data In ORE (<2GB)

Datasets can be deposited to ORE via Symplectic using the same procedure as the one used to deposit journal articles – simply select the type “dataset” rather than “journal article”

<https://researchpubs.exeter.ac.uk/>



Datasets <2GB in size can be deposited in ORE via Symplectic in the same manner that is used for journal articles. Note that the dataset and associated journal article require 2 separate Symplectic records.

# Depositing Data In ORE (≥2GB)

The ORE Deposit Tool can be used to deposit large datasets into ORE

<https://ore.exeter.ac.uk/oredeposit/>

Uses Globus Connect to transfer the data in the background without having to store the data on an intermediate server



UNIVERSITY OF EXETER Profile: Christopher Tibbs | Logout | My ORE | Advanced Search | Submit | Help | Send Feedback

Open Research Exeter (ORE)

Welcome to Open Research Exeter

As this is the first time you have chosen to make a submission through ORE there are a number of things you should be aware of:

- Ensure your data is ready for submission**  
You should take time to prepare your data as a set of files or folders in one place though this does not necessarily need to be on the University network. You should collect all your files and subfolders under one top level folder. ORE does not support the concept of folder structure so any subfolders are converted to <code>filename\zip</code> files upon submission.
- How your data will appear after submission to ORE**  
When accessing submitted items ORE recognises most common file types e.g. doc, xls, pdf, txt and will allow these to be opened directly with the related application from or within a browser. Where a file type is not recognised the file is simply downloadable. It is possible to preview the contents of zip files but not the content of the individual files within the zip.
- ORE uses Globus Online for file transfer**  
Globus Online provides a reliable way of transferring data of any size from any location without the need to keep your browser open. You will be prompted to choose University of Exeter as your identity provider and which will take you to our SSO system to login if needed. The first time you access the service you will also be asked to consent to the Exeter provider accessing your Globus related resources.
- First Time Use of the ORE Deposit Tool**  
If you wish to deposit files from a personal location like your desktop PC or laptop you will need to go through a short process to install the Globus client software and create a personal Globus "endpoint" on your PC or laptop. You need a Globus Account to create an endpoint. If the files are located in a central location like your University U. Drive or on an HPC Cluster your files will most likely be already accessible from Globus endpoints preconfigured by the University.
- Submission Tool FAQ**  
The Submission Tool FAQ provides more detailed answers to common questions on the submission tool including detailed instructions on the Globus account creation procedure and Globus endpoint creation.

Start

Submit

Start a submission  
E-DS Item Description  
Deposit Files  
Confirm and Submit  
Create a Globus Endpoint

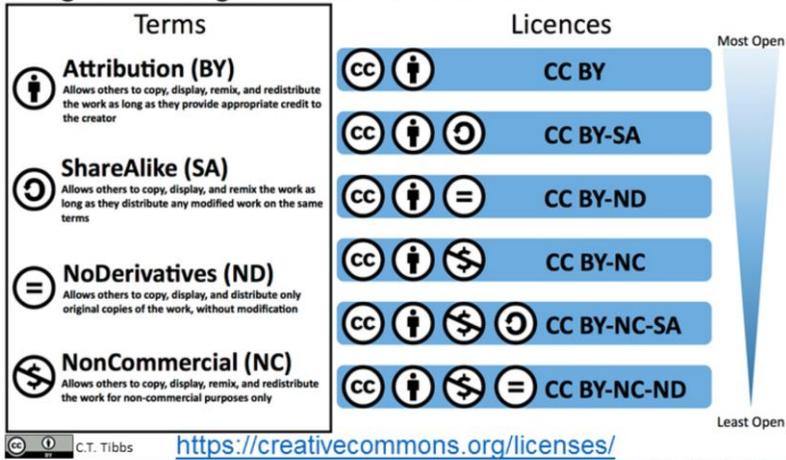
Local Links

About ORE  
How to deposit in ORE  
ORE Policies  
ORE Takehome Policy  
ORE Deposit Agreement  
E-Thesis Guides  
Open Access Information  
Research Toolkit  
Library home page  
FAQs  
Submission Tool FAQs  
Help with Creating an Endpoint

The ORE Deposit Tool can deposit large datasets to ORE. If you wish to use the tool for the first time, please contact us.

# Licensing Your Data

Licences to share, remix, and distribute legally, while ensuring authors get credit for their work



If you want to share your data and allow others to use it, you need to license your data. A licence determines how others can use, modify, and distribute your data.

# Digital Object Identifiers (DOIs)

To effectively share research data, it must be uniquely identified with a persistent identifier

All datasets deposited in ORE are allocated a DataCite DOI



University of Exeter DOIs are of the format:

<https://doi.org/10.24378/exe.XXXXX>

The screenshot shows the DataCite Search interface. At the top, it says 'DataCite Search' and 'University of Exeter'. Below that, there's a search bar with 'BL EXETER' entered. The results show '11 Works'. The first work is 'Is it safe to go back into the water? A systematic review and meta-analysis of the risk of acquiring infections from recreational exposure to seawater (dataset)'. It lists authors: Anne Frances Clare Leonard, Andrew Singer, Oboha Ukoumunne, William Gaze & Ruth Garada. The second work is 'Enhancing the magneto-optical Kerr effect through the use of a plasmonic antenna (dataset)'. It lists the author: Tom Loughran. On the right side, there are filters for 'Resource Types' (Dataset: 7, Text: 3, Collection: 1), 'Publication Year' (2018: 4, 2017: 5, 2015: 2), and 'Registration Year' (2018: 8, 2017: 3).

The University has an agreement in place with DataCite, through the British Library, and since the start of the year, all datasets deposited to ORE are allocated a DOI.

## Data Access Statements

---

All research publications **should** include a statement on how the underlying data can be accessed (a "data access statement").

- **How the data can be accessed?**
  - Web link
  - Digital Object Identifier (DOI) or other persistent identifier
- **Who must be contacted to request access?**
  - Departmental/group email address (not a personal email address)
- **On what terms are the data available?**
  - a general licence
  - a data sharing agreement must be entered into before access to the data is granted



A data access statement is a short statement that should describe where and how the underlying research data can be accessed, ideally including a link to the data using a persistent identifier. It should also explain on what terms the data are available.

## Example Data Access Statements

---

### **No new data generated**

e.g., “This study did not generate any new data.”

### **Openly available data in a repository**

e.g., “The research data supporting this publication are openly available from the University of Exeter’s institutional repository at: <https://doi.org/10.24378/exe.XXXXX>”

### **Sensitive data with restricted access**

e.g., “Due to ethical concerns, the research data supporting this publication can only be made available to bona fide researchers subject to a data access agreement. Details of how to request access are available from the University of Exeter’s institutional repository at: <https://doi.org/10.24378/exe.XXXXX>”



## Practical Steps...

---

1. Write/update a data management plan
2. Deposit research data in a repository
3. Include a data access statement in publications



## Need to know more?



Contact:

Chris Tibbs – Research Data Officer

Pat Liebetrau – Research Support Manager



[rdm@exeter.ac.uk](mailto:rdm@exeter.ac.uk)

<http://www.exeter.ac.uk/research/researchdatamanagement/>

