

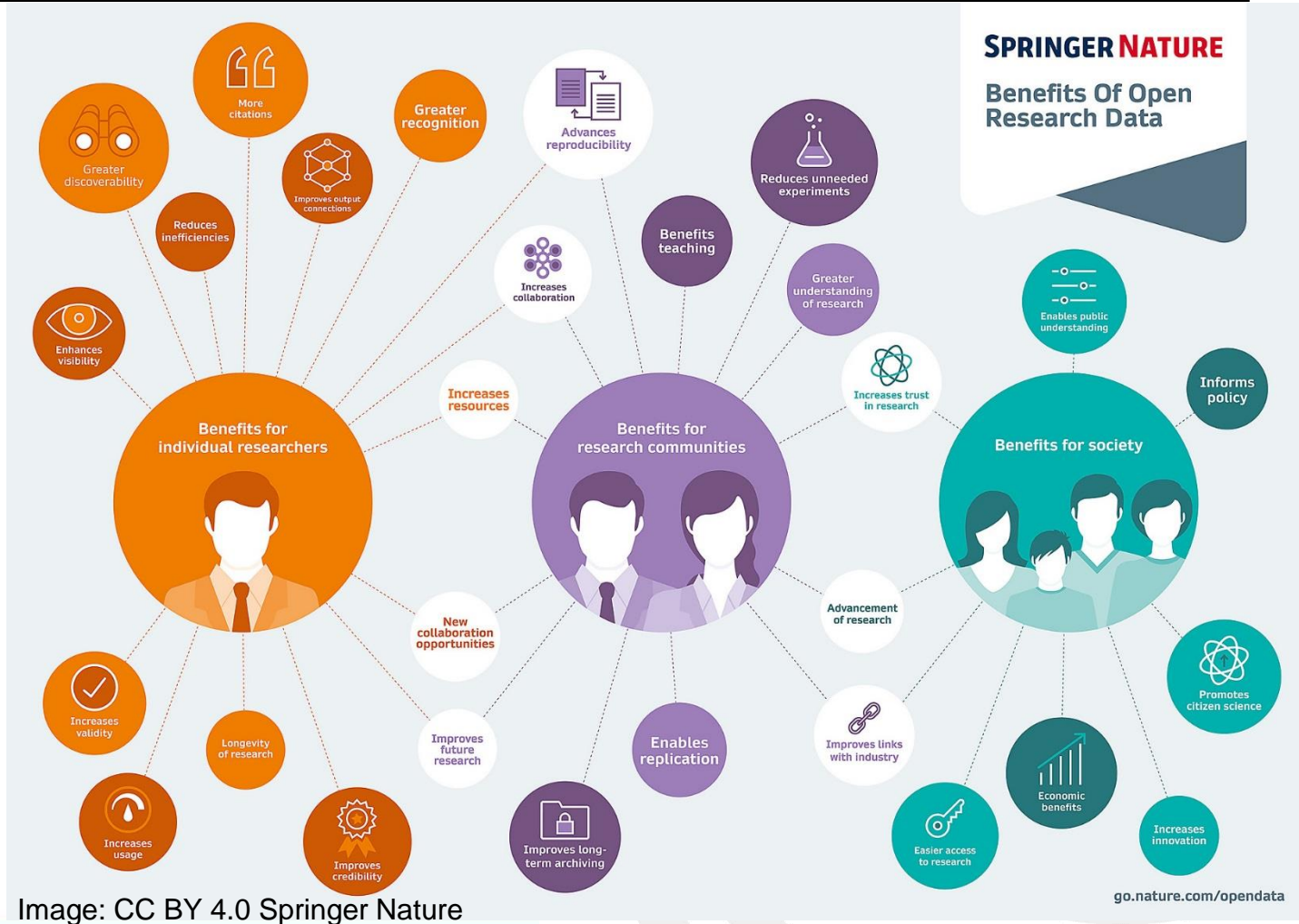
How Do I Manage My Research Data?

Dr. Chris Tibbs
Research Data Officer
rdm@exeter.ac.uk



Open Data

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



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 and sharing data

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

Exeter 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/>

- 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.

<https://www.ukri.org/funding/information-for-award-holders/data-policy/common-principles-on-data-policy/>

Data Management Planning

A decorative graphic consisting of several concentric, light gray circles of varying thicknesses, positioned on the right side of the slide.

Research Data Lifecycle

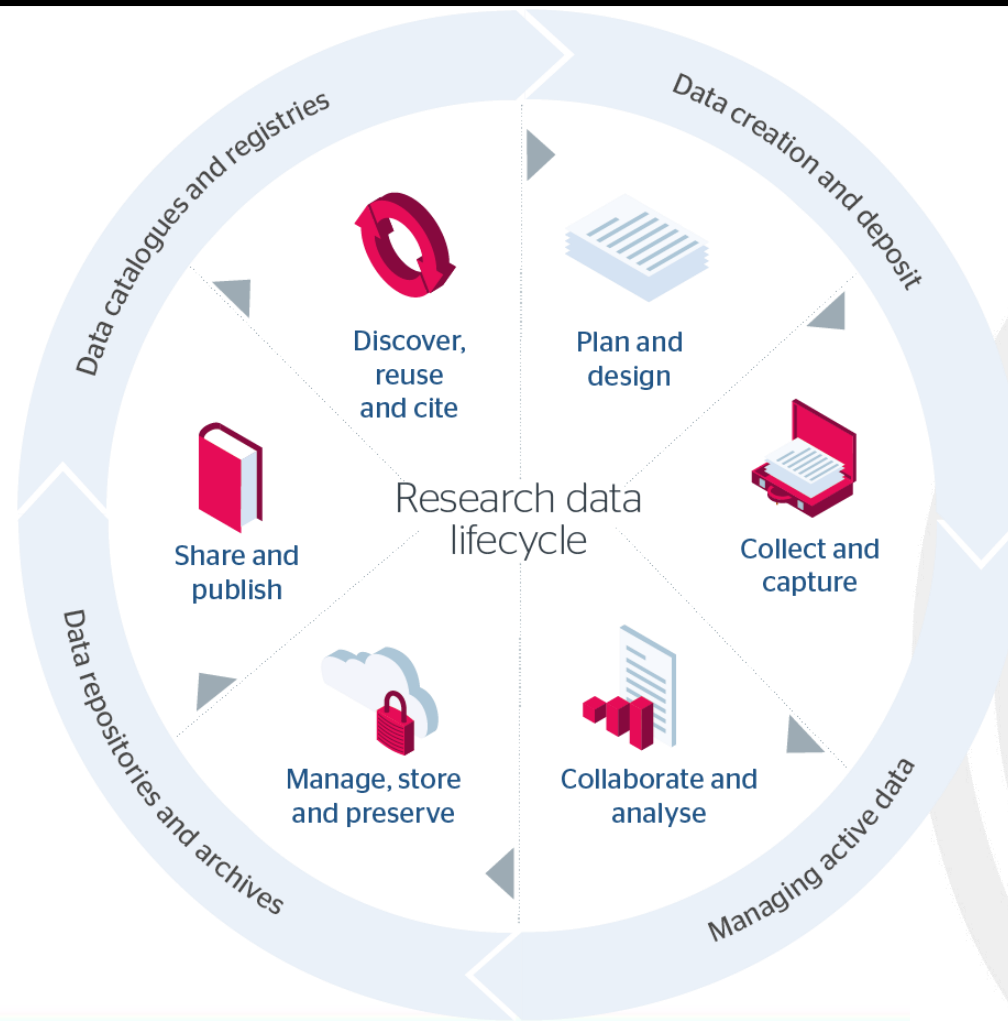


Image: CC BY-ND 4.0 JISC

What is a Data Management Plan?

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?

DMP Online



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

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



A screenshot of the DMP Online website. The top navigation bar is dark blue with the University of Exeter logo and links for Home, Public DMPs, DMP Templates, and Help. A Language dropdown menu is in the top right. The main content area has a white background with a large orange square on the left containing the DMP Online logo. To the right, a 'Welcome' message explains that the instance is for development and requires account creation. Below this is a contact email: rdm@exeter.ac.uk. On the right side, there is a sign-in and account creation form with fields for Email and Password, a 'Remember email' checkbox, and buttons for 'Sign in' and 'Sign in with your institutional credentials'. A 'Forgot password?' link is also present.

Funders That Use DMP Online



UK Research
and Innovation



wellcometrust

EPSRC

Engineering and Physical Sciences
Research Council



E·S·R·C
ECONOMIC
& SOCIAL
RESEARCH
COUNCIL



European Research Council
Established by the European Commission



Arts & Humanities
Research Council



https://dmponline.exeter.ac.uk/public_templates

What Does DMP Online Do?

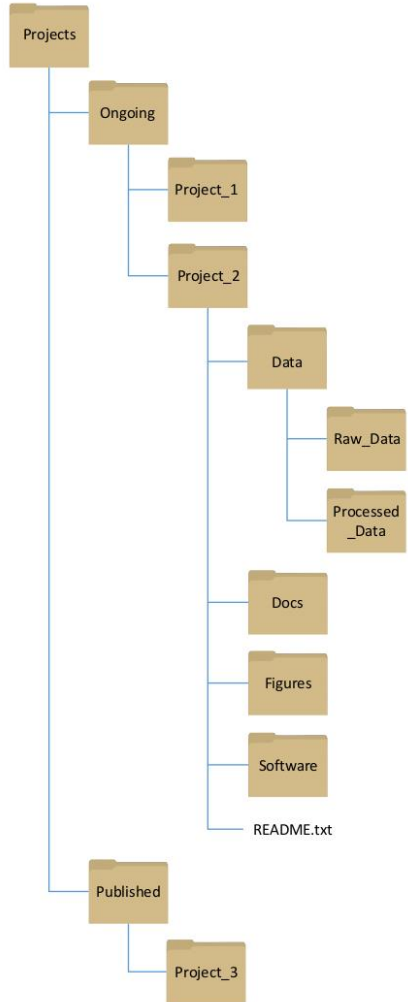
A web-based tool that enables users to

- 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. Request feedback on your DMP
- vi. Customise, export, and share DMPs in a variety of formats in order to facilitate communications within and beyond research projects

Working With Data

A decorative graphic on the right side of the slide consists of several concentric, light gray circular lines. The lines are not perfectly continuous, with some gaps, giving it a stylized, modern appearance. The circles are centered vertically and partially overlap the main title text.

Directory Structure And Filenames



Research data files and folders need to be organised in a systematic way to be identifiable and accessible to both you and colleagues and potential future users.

- Use dates in YYYYMMDD or YYYY_MM_DD format at the beginning of the file name
- Use underscores instead of blank spaces between words
- Don't use special characters: & , * % # ; * () ! @\$ ^ ~ ' { } [] ? < >
- Keep the file name as brief as possible
- Use meaningful abbreviations
- Order the elements of the file name from general to specific
- Keep the file name independent of its location
- Use appropriate file extensions

Image credit: C. T. Tibbs, University of Exeter CC BY-SA 4.0

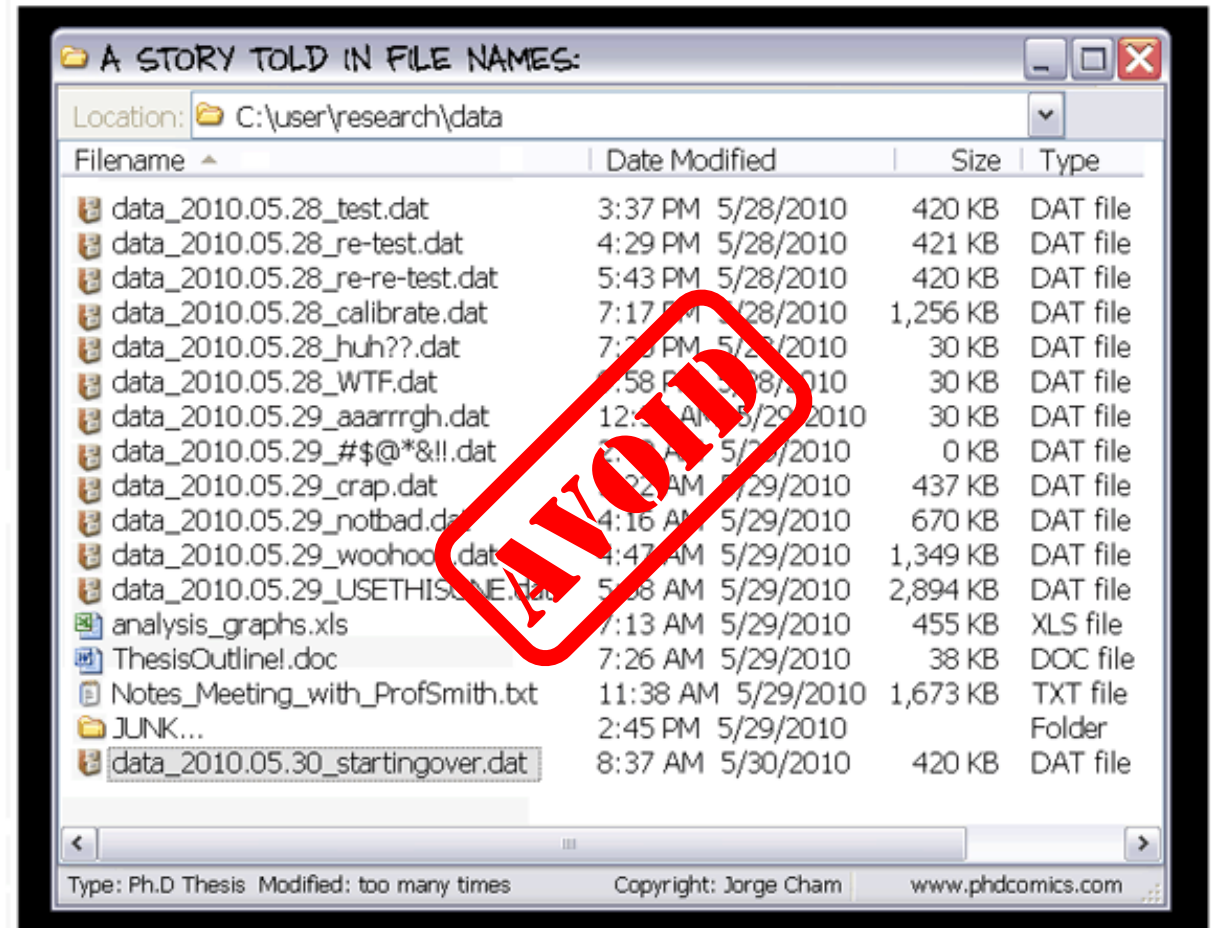
Version Control

Use a consistent method to identify and distinguish versions of data files.

This is extremely useful if you need to go back and find data from an earlier phase of the project.

Suggested strategies:

- use a sequential numbered system e.g., v1, v2
- Use version control software e.g., SVN, CVS, Git



"Piled Higher and Deeper" by Jorge Cham www.phdcomics.com

Documenting Data

Data documentation includes:

- A basic description of the project
- A description of the methodologies, protocols, sampling techniques used to produce the data
- An explanation of the content of all of the data files and how they are related to each other
- A list of any software that is required to access/analyse the data
- A detailed list of all of the variables

Electronic notebooks are:

- Searchable by any word, while physical notebooks are usually only searchable by date
- Able to support collaboration from multiple researchers
- Easier to backup and share notes with collaborators
- Accessible anywhere on multiple devices.
- Convenient to transport

Storing And Protecting Your Data

Where to store your data?

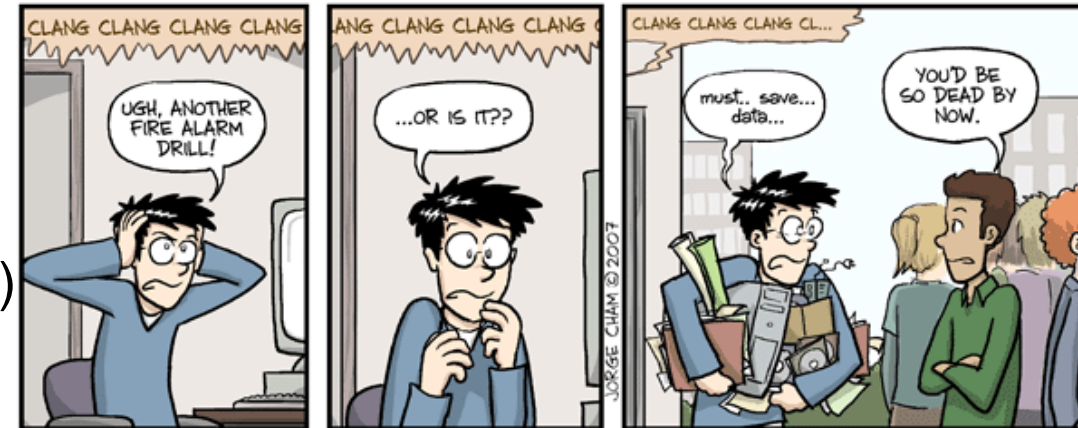
- University U:Drive
- University OneDrive for Business
- University shared network drive
- Portal storage (external hard drives, memory sticks)

Data backup strategy

- Ensure that you have at least 3 copies of your data
- Store these copies on at least 2 different media
- Keep at least 1 of these in a physically different location

Data security

- Access controls
- Encryption



“Piled Higher and Deeper” by Jorge Cham www.phdcomics.com

Preserving And Sharing Data

A decorative graphic consisting of several concentric, light gray circular arcs of varying thicknesses, positioned on the right side of the slide.

FAIR Data

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

F
Findable

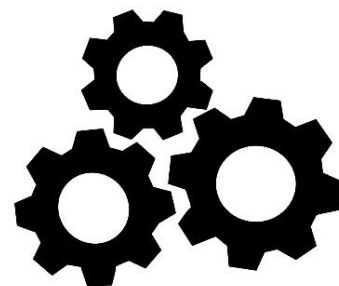


Image: CC BY-SA 4.0 Sangya Pundir

A
Accessible



I
Interoperable



R
Reusable



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

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/>

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

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

University home Studying Research Business Working here Alumni and supporters Our departments Visiting us About us

UNIVERSITY OF EXETER

ORE Open Research Exeter

Search ORE...

Advanced Search >

Home Browse Information Statistics

University of Exeter Repository

Upload Now

Business School

College of Engineering, Mathematics and Physical Sciences

College of Humanities

College of Life and Environmental Sciences

College of Medicine and Health

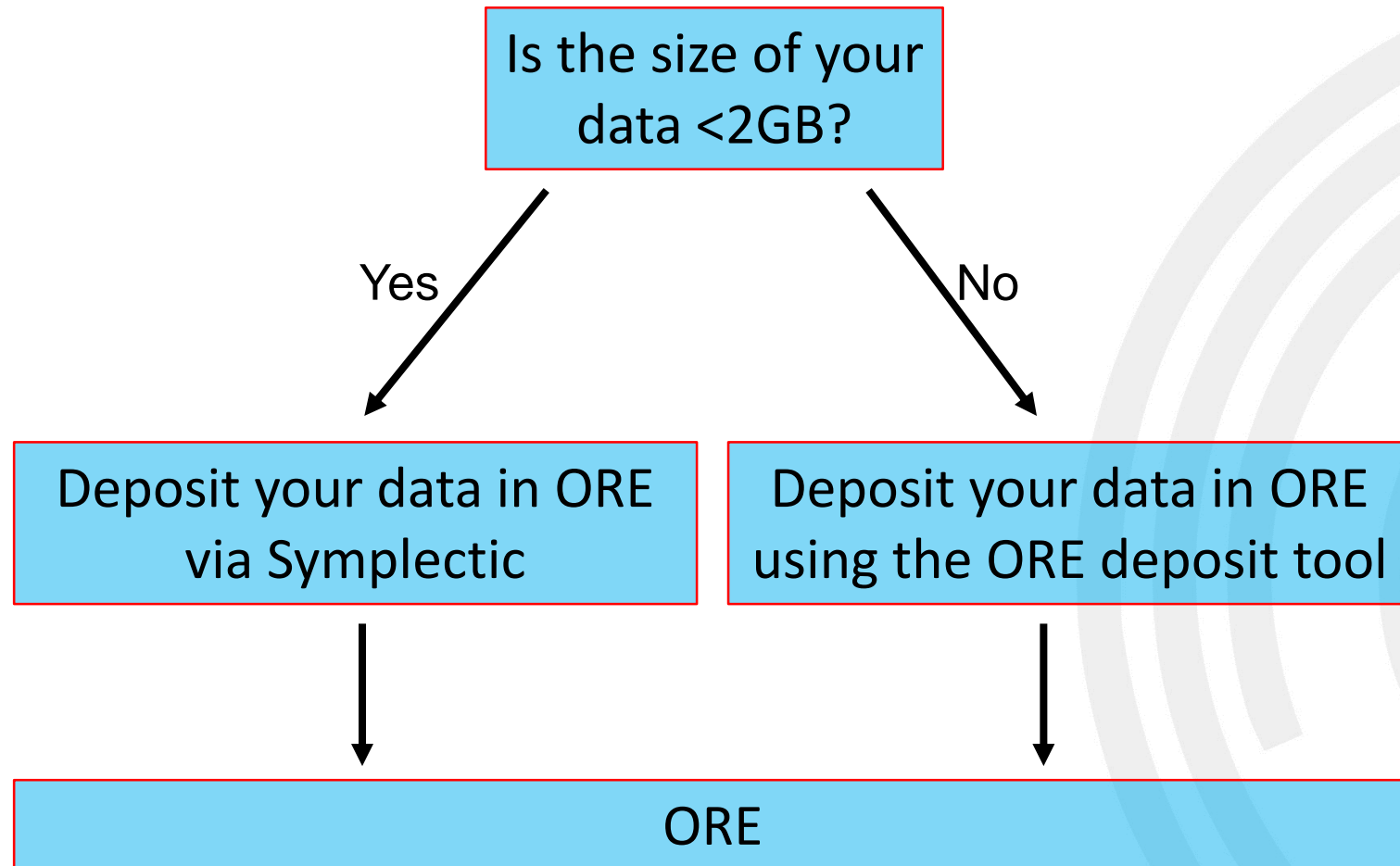
College of Social Sciences and International Studies

Doctoral College

Other Collections

Special Collections

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/>



A screenshot of the University of Exeter's researchpubs interface. The user is logged in as Christopher Tibbs. A modal window titled "Add a new..." is open, displaying a grid of publication types. The "dataset" option is circled in blue. Other options include book, chapter, journal article, conference, patent, working paper, software / code, report, performance, composition, design, artefact, exhibition, internet publication, scholarly edition, poster, thesis / dissertation, other, figure, fileset, media, and presentation. Below the grid, there is a link to "perform a bulk import of publications from your local system." The background shows the user's profile, a "Welcome" message, and various statistics like "0 publications", "0 records of impact", and "0 grants".

Depositing Data In ORE ($\geq 2\text{GB}$)

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

DSpace Home > Welcome to Open Research Exeter

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 <foldername>.zip 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.

Submit

- Start a submission
- Edit Item Description
- Select Files
- Confirm and Submit
- Create a Globus Endpoint

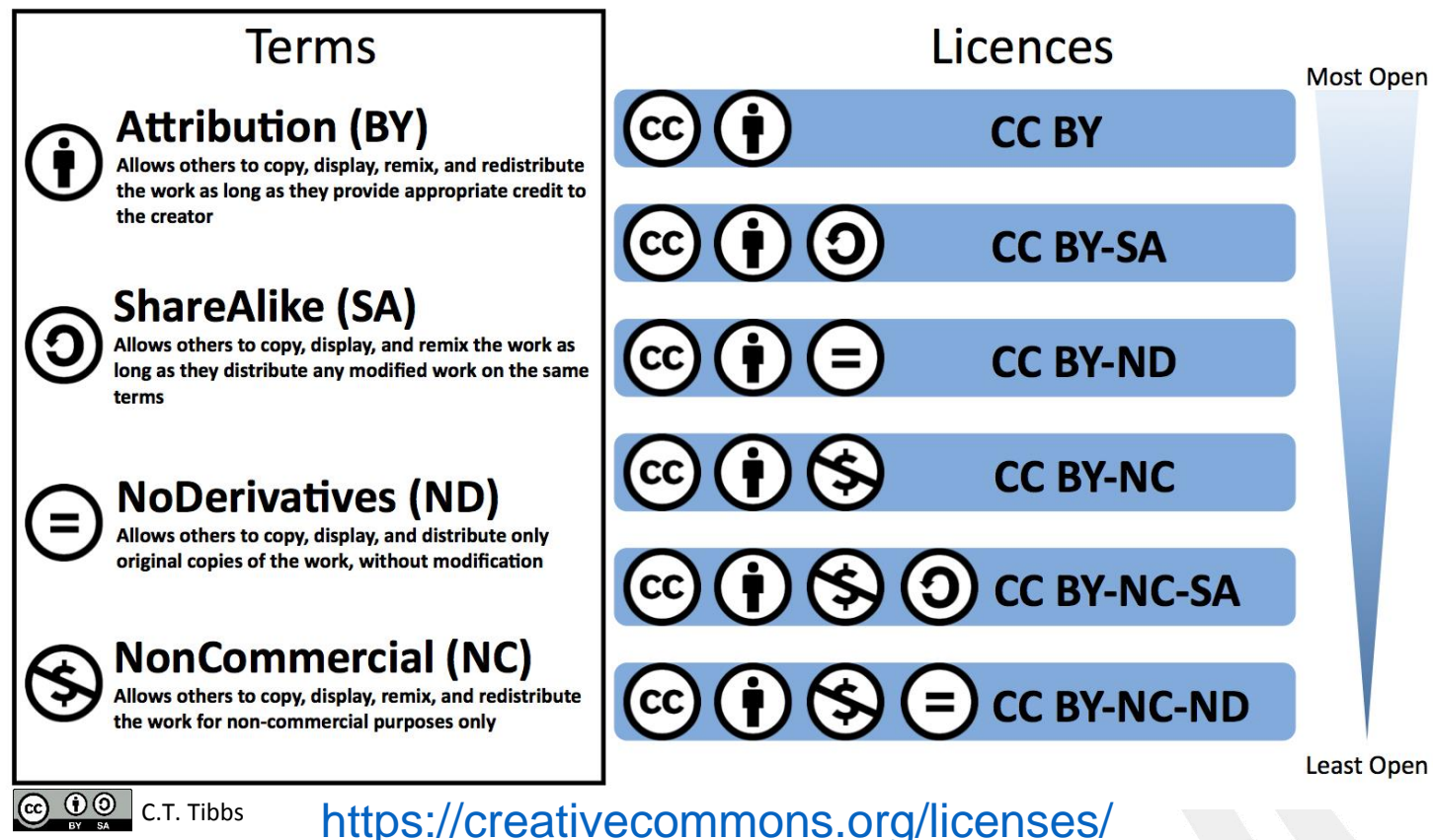
Local Links

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

Start

Licensing Your Data

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



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>

A screenshot of the DataCite Search website. The page title is "DataCite Search" and the URL is "https://search.datacite.org/". The user is logged in as "Christopher T. Tibbs". The search results are for the "University of Exeter" and show a list of 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)" by Anne Frances Clare Leonard, Andrew Singer, Obioha Ukoumunne, William Gaze & Ruth Garside. The second work is "Enhancing the magneto-optical Kerr effect through the use of a plasmonic antenna (dataset)" by Tom Loughran. The interface includes a search bar, filters for resource types and publication years, and options to cite or add to ORCID records.

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

Example Data Access Statements

No new data generated

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

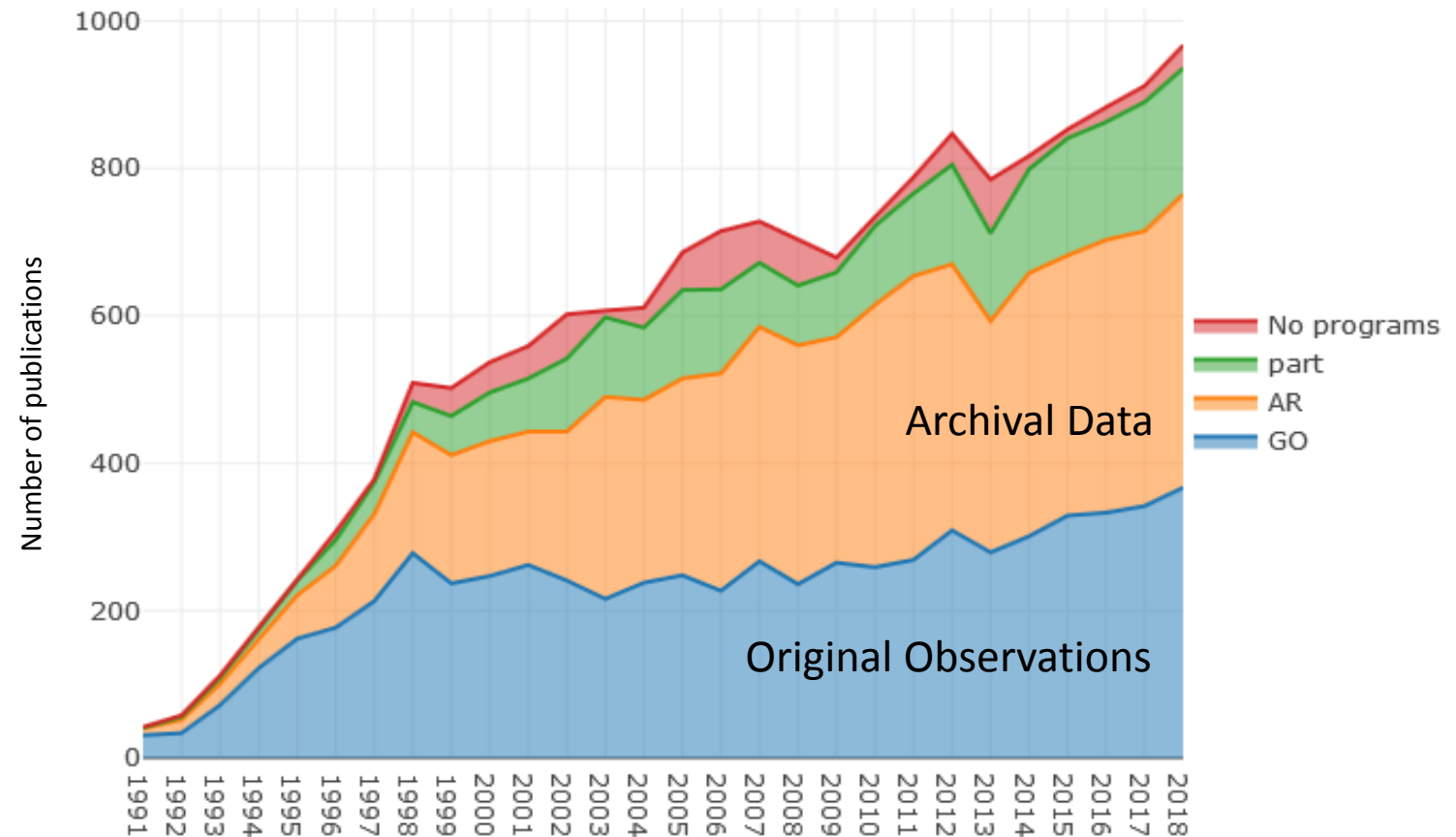
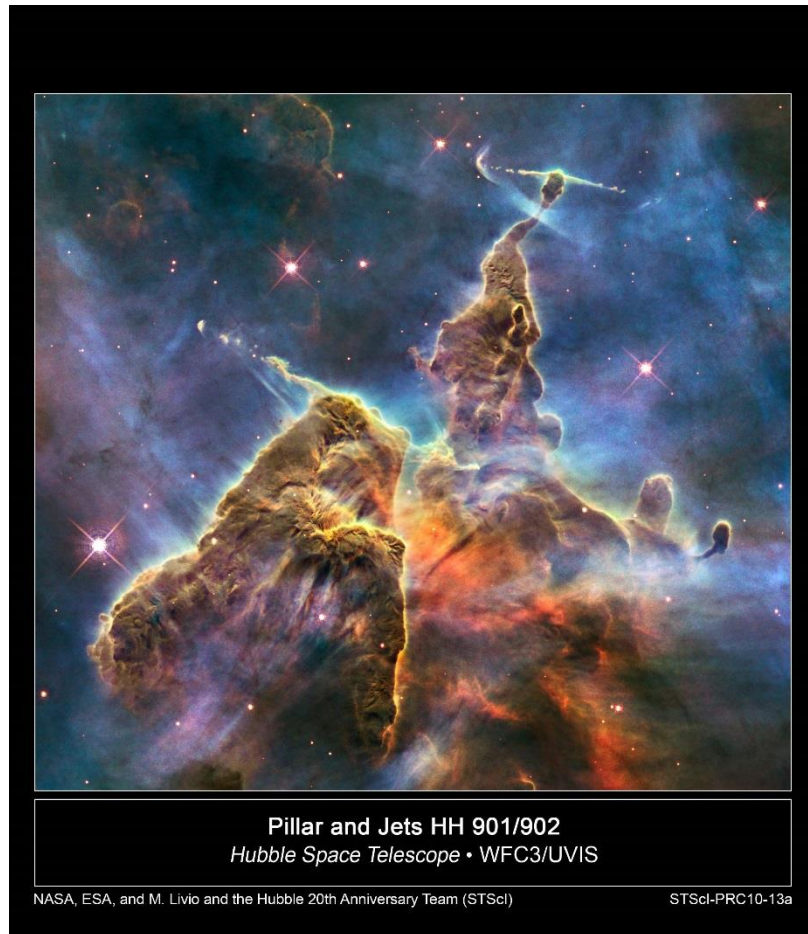
Openly available data in a repository

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>”

Why Archive Research Data?



<https://archive.stsci.edu/hst/bibliography/pubstat.html>

Further help?

Contact:

rdm@exeter.ac.uk

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

