

Research Data Management

FINISHING TOUCHES

Research Data Management

Session Aims

In this module, we'll discuss how best to finish off your project:

- Choosing the publication medium
- Sorting your data for publication
- Anonymising your data
- Choosing an archive for preservation

Research Data Management

PUBLICATION

Research Data Management Publication

Where should you publish:

- Academic journal,
- Conference proceedings,
- Community magazine,
- Newspaper.

It's important to consider

- Language,
- Impact,
- Audience,
- Cost of publication
- Their open access policy.

Research Data Management

Academic Journal

Journal Impact: How frequently the 'average article' is cited.

Calculated by comparing the number of publications in a year to the number of times those papers were cited the following year.

Science Watch provides ranking and impact factors available per subject, country, author, and institute. These are the top three journals – all science publications.

Rank	2003 Impact Factor	Impact 2000-04	Impact 1981-2004
1	Nature (30.98)	Nature (50.99)	Nature (138.97)
2	Science (29.16)	Science (50.13)	Science (138.57)
3	PNAS (10.27)	PNAS (19.11)	PNAS (72.73)



Research Data Management

Academic Journal

Journal Expenses;

Journals can charge per page, for coloured images/plots, or just a one off cost with free online additional material.

Some are free but charge subscriptions to view the articles which may affect your audience.

Your grant will include a budget for publication so balance impact with the cost of publication.

Choose your journal before getting funding to ensure your costs are covered.



Research Data Management Academic Journal

Accessibility;

Does the journal allow open access to the abstract or the whole text, limitation could damage your citation rate.

Are you prevented from submitting a manuscript to an open access publication archive?

Most journals prevent reuse of the typeset, proof-read journal paper, but you may be able to make a preprint available.

How long is access limited?

Some journals open up access after a few months of access limited to subscribers.



Research Data Management Conference Proceedings

Presentations

- Increase the audience of your research
- Allow Q&A on your work which encourage collaboration

Conference proceedings

- These articles are not peer-reviewed – arguably less reliable than journal publications
- Short articles - only a few, small pages long - so they contain very few details

It is therefore good practice to submit a journal article that contains the full details of your research in addition to a conference proceedings article.

Research Data Management Press Release

Ground breaking and of interest to the public?

Releasing details of your work in a press release increases the audience of your work.

We have a team at UH committed to helping you communicate with the media, who you can contact through the Press Office.



ReDNet, the prevention project on new drugs, scoops top European Health Award

The prestigious European Health Award 2013 has been awarded to Dr Ornella Corazza for the Recreational Drugs European Network (ReDNet), a collaborative prevention project on new psychoactive drugs led by University of Hertfordshire researchers Professor Fabrizio Schifano and Dr Corazza.

3 October 2013

Research Data Management

PRESERVING DATA

Research Data Management

Why Preserve Data?

More data available increases research productivity. You cannot reduce the time it takes to collect new data, but if the data from previous studies were available, you may not need to gather new data at all.

Increases the impact of your research as you will be cited by everyone who reuses your data.

Raise your research profile which may affect your employability and the **success of future research** funding bids.

You can also rest assured that your **data will be safe** in the long term so your future self can also reuse your data.

Research Data Management

Preserving Data

NOT EVERYTHING

Select data that is

- difficult to reproduce, if it is even reproducible,
- tools that took months, perhaps years, to produce,
- any supporting data.

What data and tools would have saved you time if they were available to use when you started your project?

What do you need to redo your work and/or expand on it?

Research Data Management

Preserving Data

It is important to retain quality data and tools of independent merit.

- your raw data,
- your reduced data, if it is at risk from software obsolescence,
- tools and scripts,
- design files,
- finalised measurements,
- supporting documentation and metadata.

Make sure there's enough information in your publication to reuse these data and therefore verify your work and use it for a concept study or comparison set.

Research Data Management

Preserving Data

Data in support...

- The data used to make the plots
- Full catalogues in digital form for easy reuse – abstracts of the catalogue may be printed, but excessive data sets are expensive to print in full.
- Electronic versions of images which contain additional 3D information.
- Videos of models – compressed to limit file size but retaining the essence.

Research Data Management

Preserving Data

Funders may require;

- Raw data is deposited as it is collected
- Reduced, cleaned, processed data is deposited with metadata for reuse
- Deposit into a specific archive

Check the requirements of your funding body regarding preserving your data!

Research Data Management

ANONYMISATION

Research Data Management

Sensitive Data

Sensitive Data is information covering:

- The racial or ethnic origin of the Data Subject
- Political opinions
- Religious or other beliefs of a similar nature
- Membership of trade unions
- Physical or mental health or condition
- Sexual Life
- The commission of any offence or criminal records

Sensitive data must be collected using an opt-in and should be carefully handled. Other classes of data which might be regarded as sensitive are data relating to children and financial information, and also data related to national defence.

Sensitive data also includes any information that is protected by University policy from unauthorized access. This information must be restricted to those with a legitimate business need for access.

Research Data Management

Personal and Confidential Information (PCi)

Personal data

Personal data are data which relate to a living individual who can be identified from those data or from those data and other information which is in the possession of, or is likely to come into the possession of, the data controller and includes any expression of opinion about the individual and any indication of the intentions of the data controller. This includes any other person in respect of the individual (Data Protection Act 1998).

Confidential data

Confidential data are data given in confidence or data agreed to be kept confidential, i.e. secret, between two parties, that are not in the public domain such as information on business, income, health, medical details, and political opinion.

The University UPR12 sets out the acceptable practice for managing PCi.

Research Data Management

Anonymisation

A person's identity can be disclosed from:

direct identifiers such as names, addresses, postcode information, telephone numbers or pictures

indirect identifiers which, when linked with other publicly available information sources, could identify someone, e.g. information on workplace, occupation or exceptional values of characteristics like salary or age

Direct identifiers are often collected as part of the research administration process but are usually not essential research information and can therefore easily be removed from the data.

UK Data Archive

Research Data Management

Anonymisation

- Assign each participant an anonymity code, which would be their identifier in a SPSS or Excel database.
- Only the PI can match an individual to their code and each participant can be given their own code in case they wish to withdraw from the study.

Most data are used en masse so individuals are difficult to identify, it is the grouped responses that researchers are interested in.

When publishing the data, information that identifies participants cannot be published such as name, address, telephone and email contact information. Apply this same anonymisation to the data sets, what doesn't need to be shared?

You can ask the participants for permission to publish certain disclosed information when you're ready to archive it if you did not do so when it was collected.

Research Data Management

Ethics approved sharing

Your sensitive data is most at risk when you share it between devices and people.

If you are working with the NHS, you will be provided with an encrypted email account that you need to use to contact them.

In your everyday environment, you can;

- encrypt your emails using **Outlook**; this also encrypts the attachments but you will need ensure the recipient can decrypt it.
- encrypt your data and files using **TrueCrypt** before attaching them to your email or using the file exchange.



Research Data Management

Open Access with restrictions

The term "Open Data" is unsettling and can have different connotations depending on who you talk to. Generally, it means that data is accessible by everyone. With respect to repositories, it is unreasonable and unlikely that all data is available without restrictions.

- If data is discoverable, then it is open to the public and accessible.
- It may not be *downloadable* by everyone
- You can restrict access to professionals, members of a particular group i.e. NHS workers.

Archives will offer different levels of access so your data is protected against misuse.

Research Data Management

NHS Data archives

Sensitive data has to be anonymised before being preserved in an archive; however, this is not a new concept as data must be anonymised for publication.

Databases with sensitive information tend to have access control so that data is not misused. The National Institute of Health and Care Excellence (NICE), provide an evidence search which is accessed using NHS Athens.

The University of Hertfordshire is a member of Athens, so if you are supporting NHS patients, you can get a NHS Athens account for free.

NHS approved databases are reviewed by their ethics committee and meet their requirements for the preservation of research data.

Research Data Management

ARCHIVING DATA

Research Data Management

Which Archive

Archive - a place where public records or other historical documents are kept. An extensive record or collection of data.

Repository - a receptacle or place where things are deposited, stored, or offered for sale

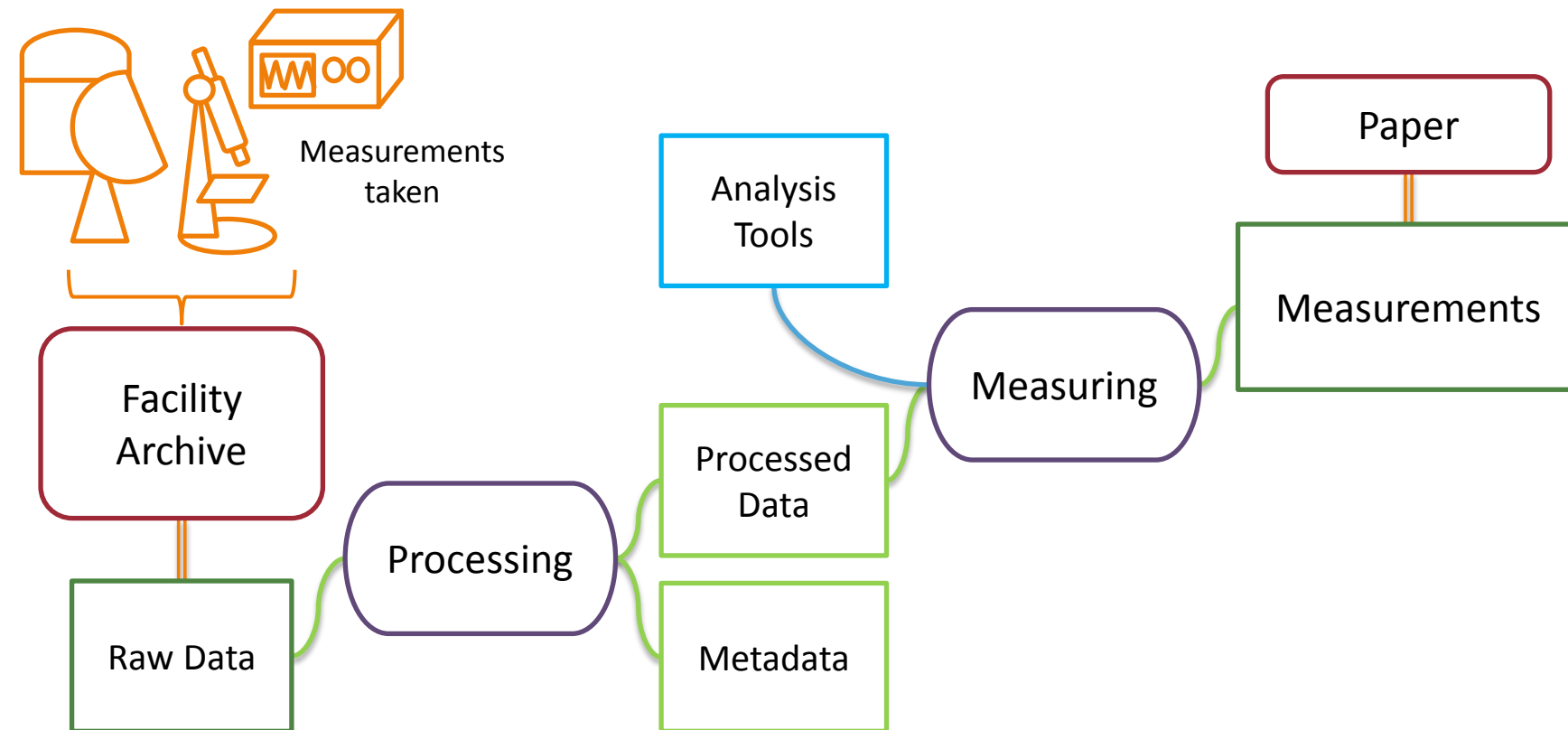
Catalogue - a list or record, as of items for sale or courses at a university, systematically arranged and often including descriptive material

Library - a collection of any materials for study and enjoyment, as films, musical recordings, or maps.

In the context of research data, repository, library, and archive are generally interchangeable. It is a place where data is deposited, managed and kept in perpetuity.

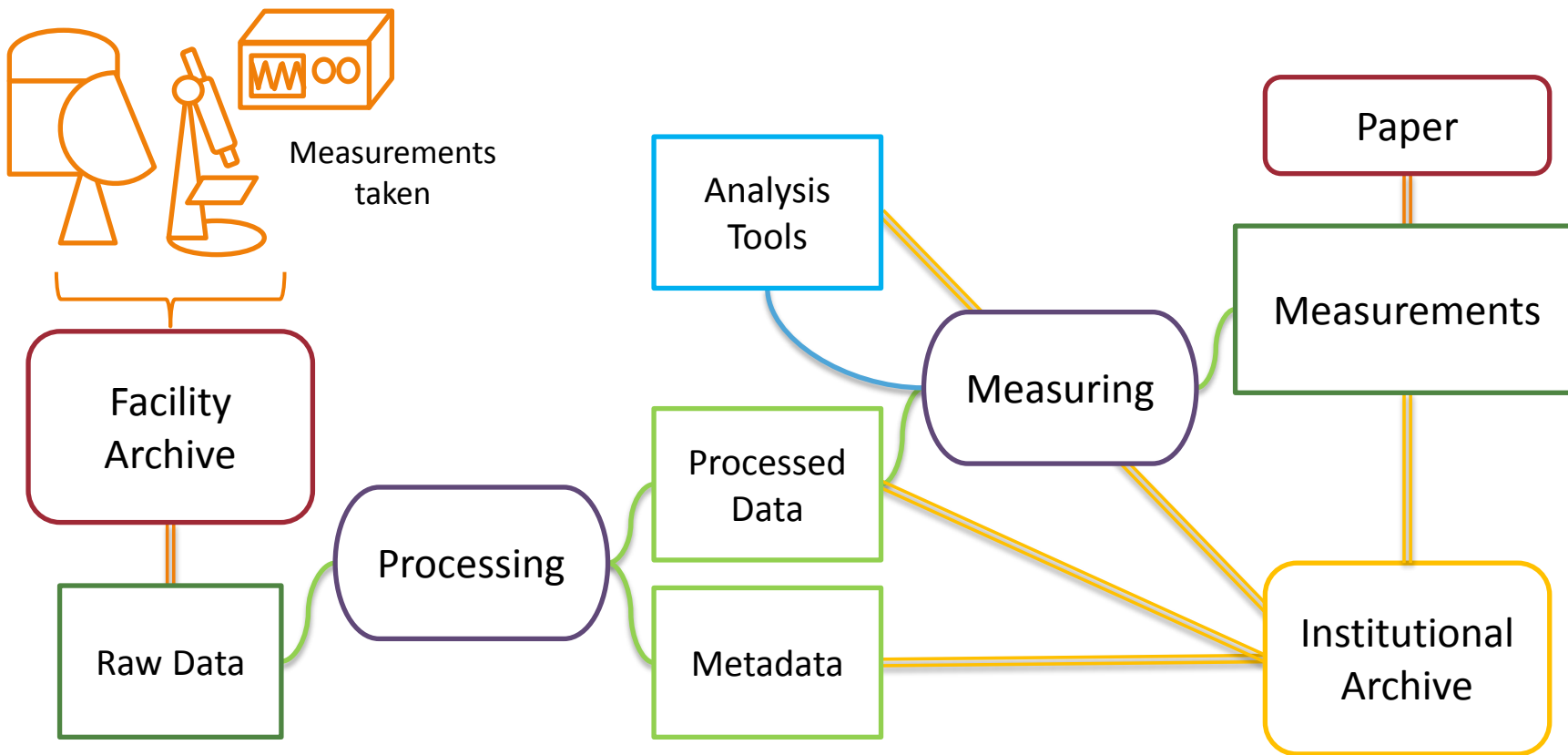
Preserving Digital Data

Standard Data Journey



Preserving Digital Data

Preservation of Data



Research Data Management Subject Archives

OpenDOAR is a authoritative directory of academic open access repositories; repositories and their content are fully searchable.

<http://www.opendoar.org>

Refine your search to subjects, languages, countries and the type of repository...

OpenDOAR

Directory of Open Access Repositories

[Home](#) | [Find](#) | [Suggest](#) | [Tools](#) | [FAQ](#) | [About](#) | [Contact Us](#)

Search or Browse for Repositories

[Recent Additions](#) [RSS1 Feed](#)

Any Subject Area Any Content Type Any Repository Type

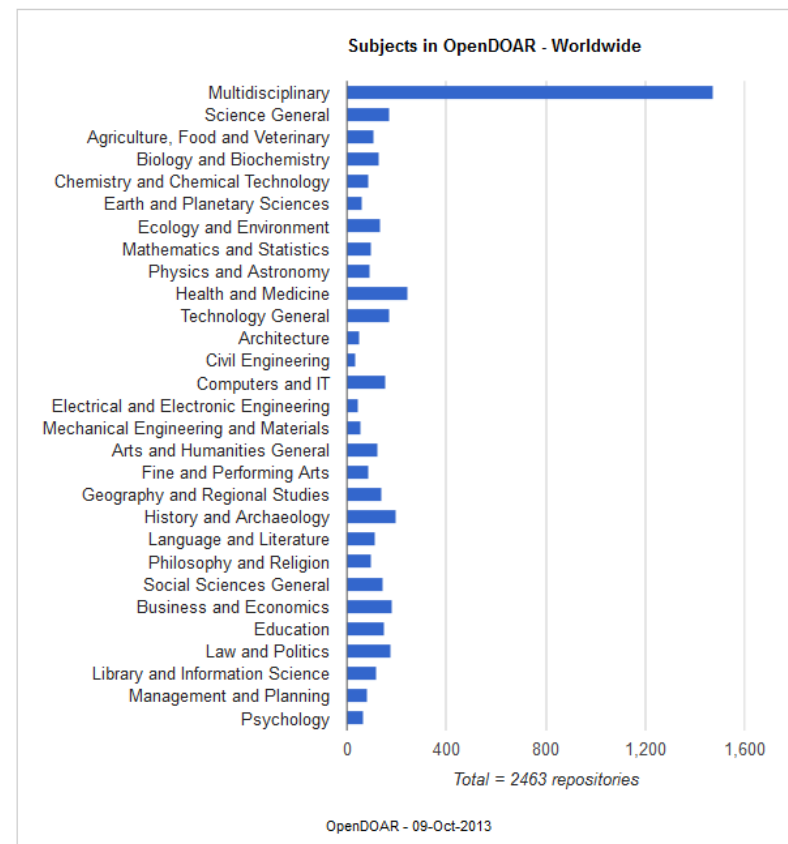
Any Country Any Language Any Software

Summaries 20 per page. Sort by: Repository Name

To search the *contents* of the repositories listed in OpenDOAR, please see our [Content Search](#) page.

Results 1 - 20 of 2463.

Page: << Previous 1 2 3 4 5 6 7 8 9 10 Next >>



Research Data Management

Institutional Archives

No dedicated subject repository?

- Deposit within an institutional repository linked to one of the investigators (PI/Col).
- Limits discoverability; subject repositories are more likely to be searched first.
- Relies on citations within publications to locate supporting and archived data.

Check if you can include data with your publication that is then hosted by the publisher.

Even if you can deposit your data in a subject repository, you may need to enter a record of these data with your institutional repository in their data catalogue.