

Research Data Management Training for Arts & Humanities

Twitter #WhiteRoseRDM

Jez Cope, Research Data Management Coordinator, Sheffield

John Lewis, Research Metadata Specialist, Sheffield

Lindsey Myers, Research Support Librarian, York

Rachel Proudfoot, Research Data Management Advisor, Leeds



Overview

- Research data management – setting the scene
- Your own research data
- Research data lifecycle
- Sharing data
- Research data – use and reuse by you
- Research data – reuse by others
- Metadata and documentation
- Research data for the long-term
- Data management planning
- Sources of help and information

Research data management: setting the scene



Research data sharing in practice

Siobhan Davies Replay

The Archive of Siobhan Davies Dance, containing thousands of fully searchable digital records (data) of choreographic work.

It includes film footage, images, audio and text.

The screenshot shows the website for the Siobhan Davies Dance Archive. The header is dark blue with the 'SIOBHAN DAVIES REPLAY' logo on the left and navigation links on the right. The main content area is white and features a search bar and a list of dance works. The 'Dance Works' list includes titles like 'Sphinx (1977)', 'Plain Song (1981)', and 'Rotor (2010)'. The 'Rotor (2010)' entry is highlighted, showing its title, date, artists, and a summary. A thumbnail image of a dance performance is visible on the right side of the page.

SIOBHAN DAVIES REPLAY

Home
Foreword
About the Archive
Guest Scrapbooks

Siobhan Davies
Dance Works
Related Projects
Dancers
Collaborators

Sign In
Register
Help and FAQs

The Archive of Siobhan Davies Dance

Advanced Search
Back to Search Results

Title: Dance Work
Work: ROTOR

Date: 2010

Artists: Siobhan Davies, Andrea Barzley, Lindsey Batcher, Anne Lok, Charlie Morrison, Massimo Barzoli, Sam Collins, E V Crowe, Angela de la Cruz, Helen Fergan, Alexandra Hughes, Alice Oswald, Claire Tunney, Sam Tyers

Summary
ROTOR is an ensemble of performances, sound, and installations created by Siobhan Davies Dance. A new dance work by Siobhan Davies was filmed from above and its patterns created a score. Distinct elements of The Score triggered new work from eight commissioned artists, which responded not only to images from the dance but also to the energies, counterpoints or characters developed by the dancers' actions. The work was initially presented at Siobhan Davies Studios, and South Jordan Gallery, London, and has subsequently been shown at The Woodruff Art Gallery, Manchester and Central Station, Edinburgh.

Dance Works

1970s
Sphinx (1977)

1980s
Plain Song (1981)
Rushes (1982)
Carnival (1982)
Bridge the Distance (1983)
The Run to Earth (1988)
And Do They Do (1988)
Enfance (1988)
Wise Man Sleeps (1988)
Wyrning (1988)
Sounding (1989)
Cover Him With Grass (1989)
Drawn Breath (1989)

View all records
Print this record
Save metadata

Related Media

- Movies
- Images
- Audio
- Text
- Object
- Profile

<https://youtu.be/mjV0nB6mnR0>

What is data management?

Data management is a general term covering how you organize, structure, store, and care for the information used or generated during a research project

It includes:

- How you deal with information on a day-to-day basis over the lifetime of a project
- What happens to data in the longer term -what you do with it after the project concludes

Your own research data



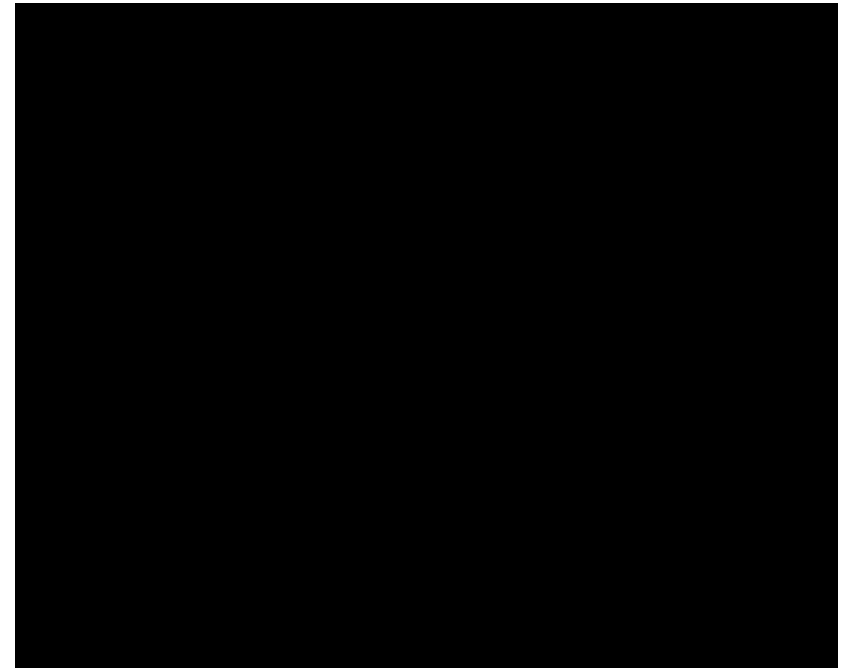
What is data?

statistics, spreadsheets, experimental results, databases
works of art, images, sketchbooks, performances, video
recordings, artefacts, photographs, interview recordings
or transcripts, surveys, logbooks, physical objects,
literary or historical works, storyboards, trials,
prototypes, ceramic glaze recipes, found objects,
correspondence, emails, journals, workbooks, music
recordings, websites, exhibitions, scripts or musical
scores ...



Creating data as historians: an example

History academics at the School of Advanced Study (University of London) talk about what research data is for historians.



<https://youtu.be/i2XSwTdY45I>

Research data is ...

“A reinterpretable representation of information in a formalized manner suitable for communication, interpretation, or processing.” Digital Curation Centre

All the information you use as an integral part of your research

- Both physical (e.g. print) and digital material is data
- Digital data can be:
 - created in a digital form ("born digital")
 - converted to a digital form (digitised)

Your research

1. What data will be produced?

- What data will be collected or created during the course of your project?
- How will you collect or create the data? What methods/standards will you use for data creation?
- If pre-existing data is being used, where will it come from? How will it be used?
- How much data do you expect to generate?

DMP template - digital copy: <http://bit.ly/1AoFs9F>

*“[...] I’m just like anyone else I’ve got **boxes of stuff**, I’ve got a garden shed and then I’ve got files, I’ve got electronic files and I’ve got physical files, I’ve got ring binders full of clippings, full of photographs, and I’ve got documents of exhibitions that I’ve been in, I’ve got catalogues of exhibitions I’ve been to [...].”*

Artist-researcher I, Institution B

KAPTUR project: managing visual arts research data

www.vads.ac.uk/kaptur

Data appraisal: What data should you keep?

Data Types	Value	Examples
Derived or compiled data	Reproducible by examining the same sources (but this can be expensive)	Data gathered from primary and/or secondary sources, databases.
Observational data captured 'in the moment'	Usually irreplaceable	Oral history, survey results, video of performance

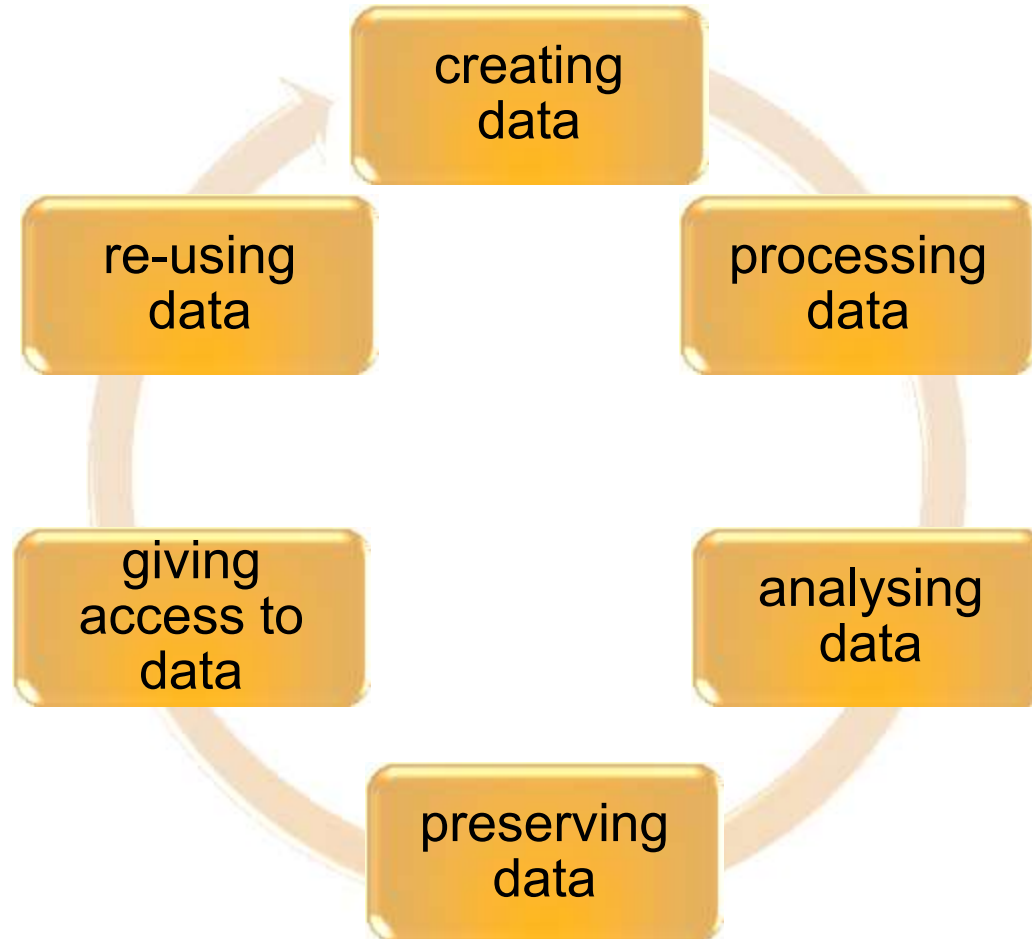
Selecting what to keep and what to bin - checklist:

<http://find.jorum.ac.uk/resources/10949/17171>

Research data lifecycle



UK Data Archive research data lifecycle



Sharing data



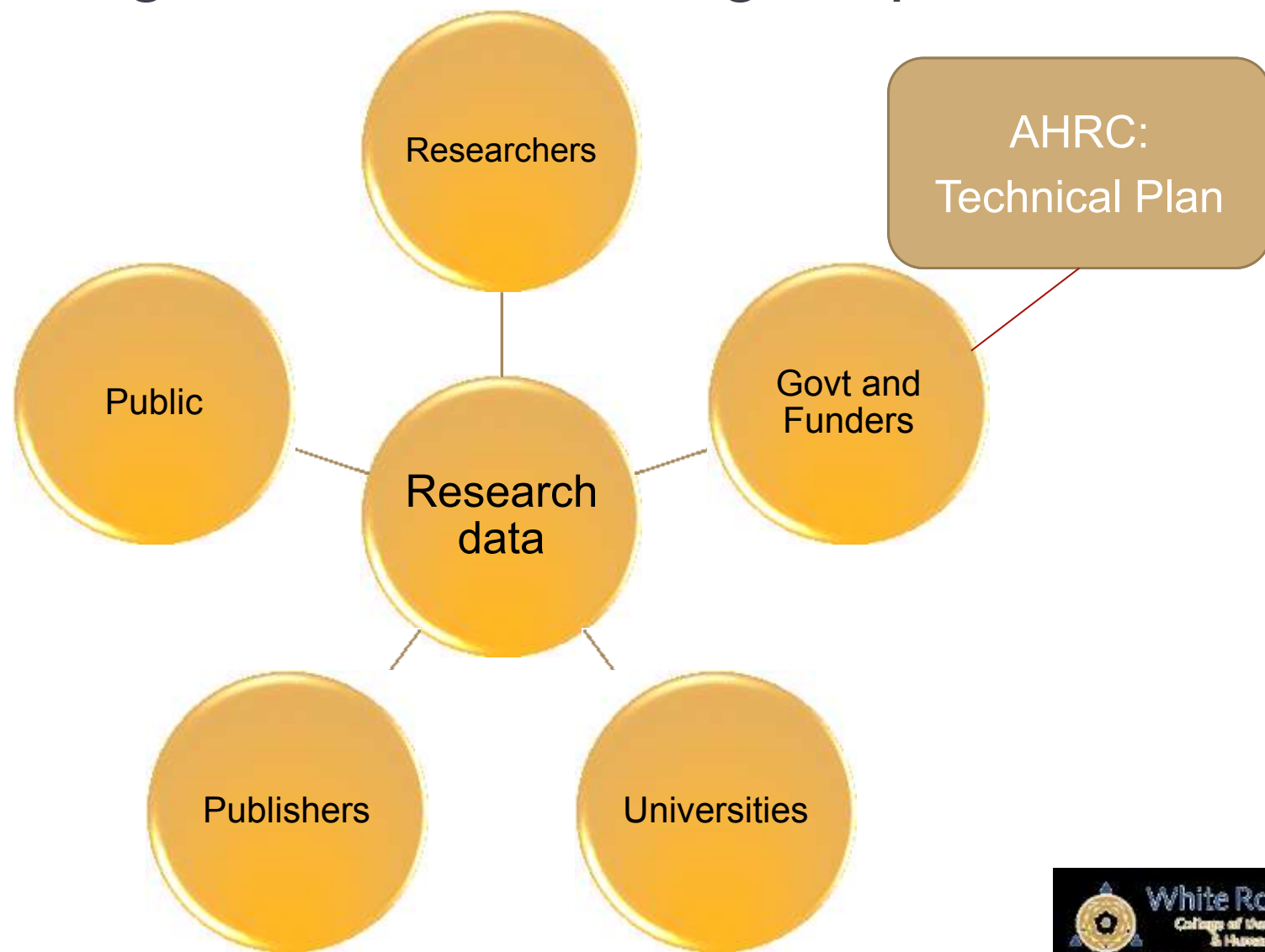
Data management and sharing: rewards

- Increase research efficiency, save time “.. as I have learned to my cost over the last 25 years..”
- Enhance data security
- Prevent duplication of effort
- Build new research networks
- Citations to your published papers may increase if you share the underlying data



(Piwowar & Vision 2013; Dorch, 2012, Wickerts et al 2011)

Data management and sharing: requirements



Data sharing: responsibilities

“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 that does not harm intellectual property.”

(RCUK Common Principles on Data Policy, www.rcuk.ac.uk/research/datapolicy/)

- **timely and responsible**

data sharing \neq open access

but

planning ahead can avoid unnecessary data
sharing restrictions

Research data use and reuse: by you



**DON'T LET THIS BE
YOU!**

Data storage

Research data must be:

- kept safe and secure, avoiding data loss
- kept in a manner that is compliant with legal and ethical obligations, and (if applicable) funder requirements.

What storage device is most appropriate to your needs?
(volume / security)

Storage – Do's

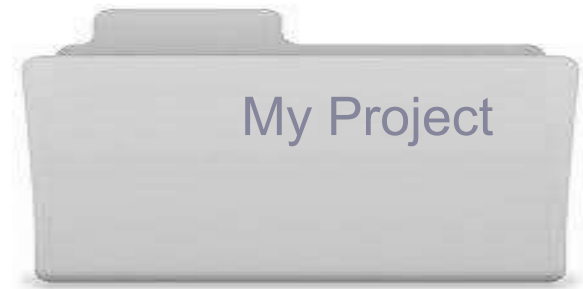


Each University offers a range of facilities to securely store your data, helping it live a long and useful life.

We recommend that digital data is stored on:

- the University's networked file servers (your individual or shared networked drive)
 - required for guaranteed UK storage
 - required for “must be kept on site”

Storage – Don'ts



Don't ...

- keep your data **just on** your working machine (laptop or a desktop) – it's the perfect way to lose your data easily and permanently
- use USB memory sticks – if you have to, encrypt them and **never keep the only copy of your data** on one
- upload personal/sensitive data to services the University does not have a contract with (e.g. Dropbox)

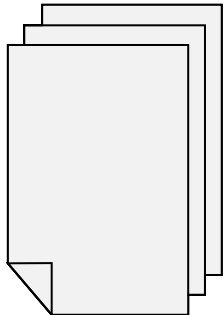
Backing up

What is the risk of losing your data?
(likelihood / consequence)



Don't have only 1 copy of your data or use only 1 type of data storage (**LOCKSS**)

multiple copies



keep in different places



automate



Your research

2. How will the data be stored and backed up during the lifetime of the project?

- Where will your data be stored?
- What sort of storage media will you use?
- How often will your data be backed up? Who is responsible for doing this? Where will your back-up copies be stored?

**Can you find what you
need, when you need it?**



Reusing data

Do you still understand your older work?

Is the file structure / file naming understandable to you and to others?

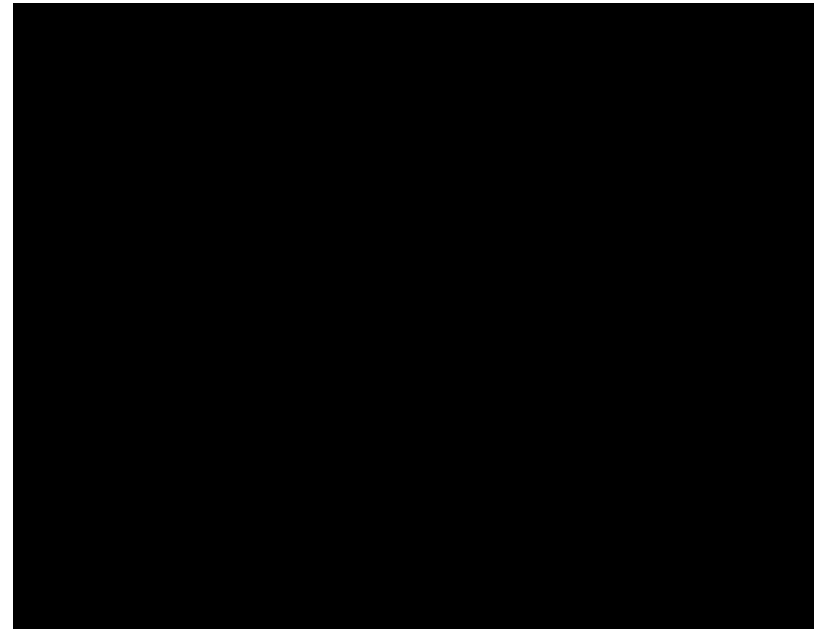
Which data have been kept?

Which data were discarded?

How much was planned, how much was circumstance?

Importance of good file management

Professor Jeff Haywood
(Vice Principal Knowledge
Management at University
of Edinburgh) discusses
the importance of good file
management in research.



<https://youtu.be/i2jcOJOFUZg>

Video credit: This video was recorded as part of
the University of Edinburgh [Research Data
Management Training \(MANTRA\) project](#).



File (and folder) naming

Decide on a file naming convention at the start of your project, one that is meaningful to you.

Useful file names:

- are consistent
- are concise (<20 characters) but informative
- classify broad file types
- allow you to find the file easily
- do not contain symbols, underscores or spaces
- should not conflict when moved from one location to another.

File formats

Use an open file format to ensure future readability

MS Word / PowerPoint with Macro	x
PDF / A	✓
MS Excel	x
Comma-separated values (.csv) / Tab-delimited (.tab)	✓
Bitmap (.bmp)	x
Photoshop (.psd)	x
JPEG-2000 (no compression)	✓
TIFF (v6 uncompressed)	✓
Real Audio (.rm)	x
Windows Media Audio (.wma)	x
FLAC (.flac)	✓
WAV (.wav)	✓

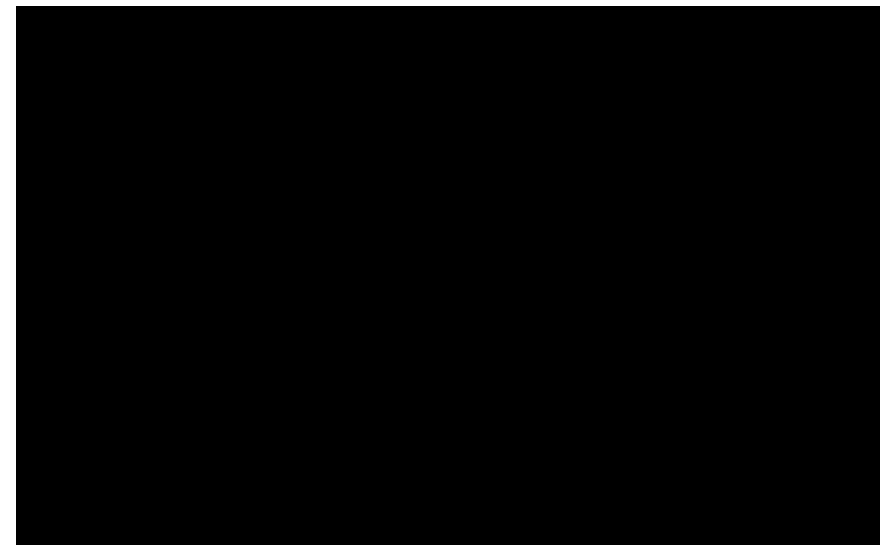
Research data reuse: by others



Reuse by others – a bad example

Judy Benign has identified some data that she might use in her research ...

This is what shouldn't happen when a researcher makes a data sharing request!



https://youtu.be/66oNv_DJuPc



Reuse by others – a bad example

This is what shouldn't happen when a researcher makes a data sharing request!

What issues (bad RDM practice) does the video highlight?

Metadata and documentation

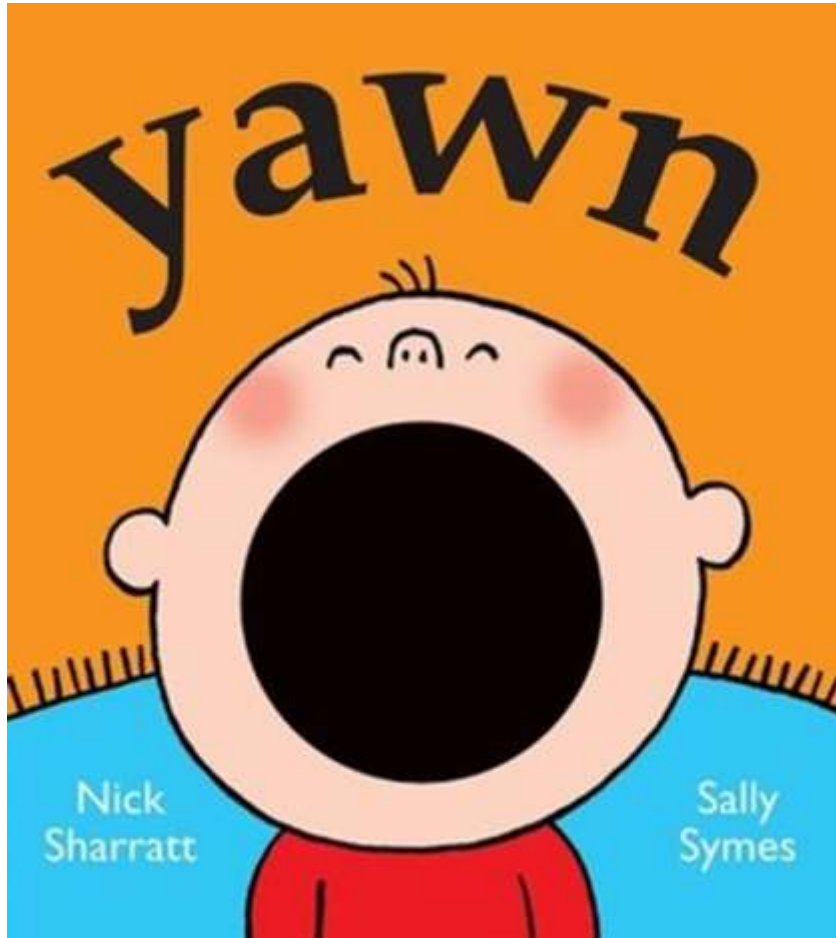


Will someone else understand your data if it isn't documented?

*"The single most useful thing you can do to ensure the long-term preservation of your data is to plan for it to be reused. **Imagining it being reused by someone else who has never met you and who never will meet you**, will cause you to approach the creation and design of your data in a new light. In short, always plan for re-use"*

Prof. Julian D. Richards, Director, Archaeology Data Service, University of York

Documentation and Metadata





Metadata is
a love note
to the
future..

Jason Scott

Metadata and Documentation

- Metadata – structured information about a resource
 - Metadata a type of documentation
 - Metadata has multiple functions
 - e.g. [Flickr](#)
 - Key functions include discovery and reuse
-
- More info: <http://www.data-archive.ac.uk/create-manage/document>

Metadata and Documentation

- Data repositories & archives require some generic descriptive metadata, e.g. author, title, publication date, Creator, Abstract, Keyword, Data type, Geographic coverage
- Digital object identifier (doi)

ADS collection level metadata (not all fields)

Title	title (and any alternatives) for the dataset.
Introduction	brief summary of project aims
Overview	summary of dataset content
Interface images	two illustrative images
Language	language of the dataset
Identifiers	associated identifiers specific to the collection
Related Resources	bibliographic references to related resources
Project Dates	when the dataset was created, or other dates as appropriate
Project Funders	external bodies and institutions that provided financial support
Copyright Holders	all copyright holders, normally employer if created during your work
Data Creator	individuals and organisations collection should be attributed to
Subject	keywords for the subject content of the dataset

Metadata for reuse

- Original creator (Judy Benign)
- Field name meanings (Judy Benign)
- Data guide / structural map
- Data format
- Research design and methodology
- Field notes
- License conditions

Practicalities

- Depends on size and complexity of data
- E.g. readme file for multiple file folders
- E.g. spreadsheet
- E.g. blog post tags

- Discipline specific metadata standards
<http://www.dcc.ac.uk/resources/metadata-standards>

Metadata about you

- ORCID (Open Researcher and Contributor ID)
- International, persistent, unique identifier for researchers
- Connects you with your outputs and activities
- Non-sensitive information: name, email, organization and research activities
- Likely to be adopted UK-wide
- Time saving: create information once, use many times
- E.g. ORCID profile <http://orcid.org/0000-0002-3510-0728>

Voluntary metadata



ZOONIVERSE

After lunch: Breakout sessions

First session: starts 1.15 pm

Second session: starts 1.45 pm

1. Copyright and research data – Beccy Shipman
2. Data interpretation - Renee Timmers
3. Preparing data for deposit - Katie Green
4. Preparing data for reuse - Tim Evans
5. Research data and ethics - Kevin Macnish

Meet back here 2.15pm – feedback from your breakout groups



Welcome back

Feedback from breakouts





Research data for the long term



Data repositories / Data centres

- Does your funder or your publisher have a policy?
- Is there a widely used data centre in your field? See re3data.org
- Subject/national data repository – [UK Data Archive](http://www.ukdataservice.ac.uk)
- Interdisciplinary (including negative data) – [FigShare](http://www.figshare.com), [Dryad](http://www.dryad.org)
- Institutional - archive.researchdata.leeds.ac.uk

Advantages: permanent / stable, findable, safe and controlled environment, citable

Your research

5. What are your plans for long-term archiving and data sharing after submission of your thesis?

- Will any of the data supporting your thesis be made available to others on request or open access (including supervisor and research group)? If so, how will this be facilitated?
- Who, if any, are the anticipated future users of any data / resources from your research?
- Are there any funding body or institutional requirements regarding reuse of, or open-access to, your data?
- Do any legal or ethical issues restrict the extent to which your data may be shared?

Open access publishing

Research literature is available online

- Via an online journal (gold OA)
- Via an institutional repository (green OA)

No cost to the reader who can, depending on licensing, re-use the information e.g.. download, copy, distribute, text and data mine without financial, legal, or technical barriers

See OA explained

<http://www.phdcomics.com/comics.php?f=1533>

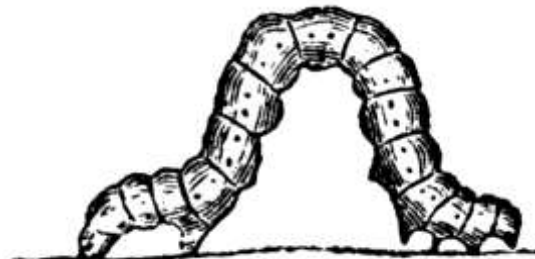
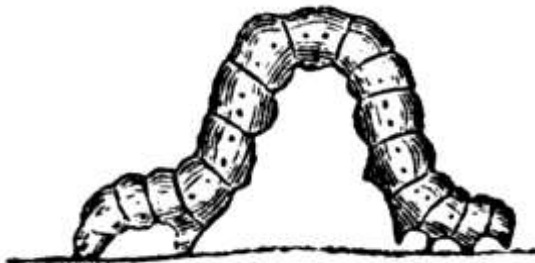
Scholarly practice

PLOS' New Data Policy: Public Access to Data

“authors must make all data publicly available, without restriction, immediately upon publication of the article”

European Commission – Horizon 2020 (€80 billion 2014-20)

“.. In addition to the announced Pilot on Open Research Data, open access to scientific publications is now mandatory for all scientific publications resulting from H2020 funded projects.”



Data Citation

- Data as a primary research output

- DataCite format

<https://www.datacite.org/whycitedata>

Creator (PublicationYear): Title. Version.
Publisher. ResourceType. Identifier

Thomson, Rachel and Kehily, Mary Jane (2011):
Dynamics of Motherhood. V.2. Timescapes Archive.
[Dataset]. <http://doi.org/10.5518/1234567>

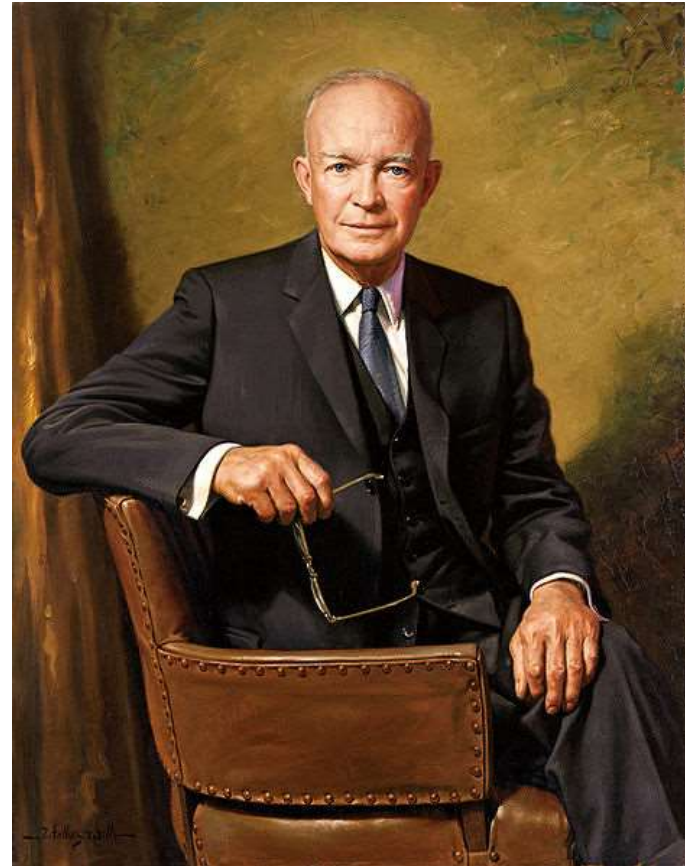
Data management planning

Good research data management
requires good planning



‘In preparing for battle,
I have always found
that plans are useless
but **planning is
indispensable.**’

Dwight D. Eisenhower



Data management plan (DMP)

A DMP is a formal document which outlines all aspects of your data - i.e. what you will do with data during and after your research project.

**‘Avoid panic and frustration’
(UKDA)**

[www.dcc.ac.uk/sites/default/files/documents/publications/reports/guides/How to Develop.pdf](http://www.dcc.ac.uk/sites/default/files/documents/publications/reports/guides/How%20to%20Develop.pdf)

University DMP requirements

Leeds

<http://library.leeds.ac.uk/research-data-policies>

“A data management plan that explicitly addresses the capture, management, integrity, confidentiality, preservation, sharing and publication of research data **must be created for each proposed research project or funding application.** Sufficient *metadata* shall also be created and stored to aid discovery and re-use.”

Sheffield

www.shef.ac.uk/ris/other/gov-ethics/grippolicy/practices/all/rdmpolicy

“All research proposals for **funded research** should include a data management plan.”

York

www.york.ac.uk/research-data-management

“producing and adhering to a Data Management Plan (DMP) as **required by the funder or specified by the Department**”

Data management plan (DMP)

Many funders require a data management plan as part of grant applications.

DCC Funders' data plan requirements

www.dcc.ac.uk/resources/data-management-plans/funders-requirements

Funder requirements vary

Common themes

- Description of data
- Data collection methods
- Ethics and IPR
- Plans for data sharing
- Strategy for long-term preservation

Format

- 1-2 sides A4 as attachment or via Je-S form
- Typically prose statement covering suggested themes
- Justify RDM decisions, including resources

AHRC DMP requirements

If digital outputs are planned, a **technical plan** should be submitted at the grant application stage. This should **not exceed 4 pages** and must use headings provided by AHRC:

- Section 1. Summary of digital outputs and digital technologies
- Section 2. Technical methodology
- Section 3. Technical support and relevant experience
- Section 4. Preservation, sustainability and use

www.ahrc.ac.uk/Funding-Opportunities/Research-funding/RFG/Application-guidance/Pages/Technical-Plan.aspx

Create a data management plan using the DMPonline tool



Welcome.

DMPonline has been developed by the **Digital Curation Centre** to help you write data management plans.

Screencast on how to use DMPonline



Sign in

[Forgot your password?](#)

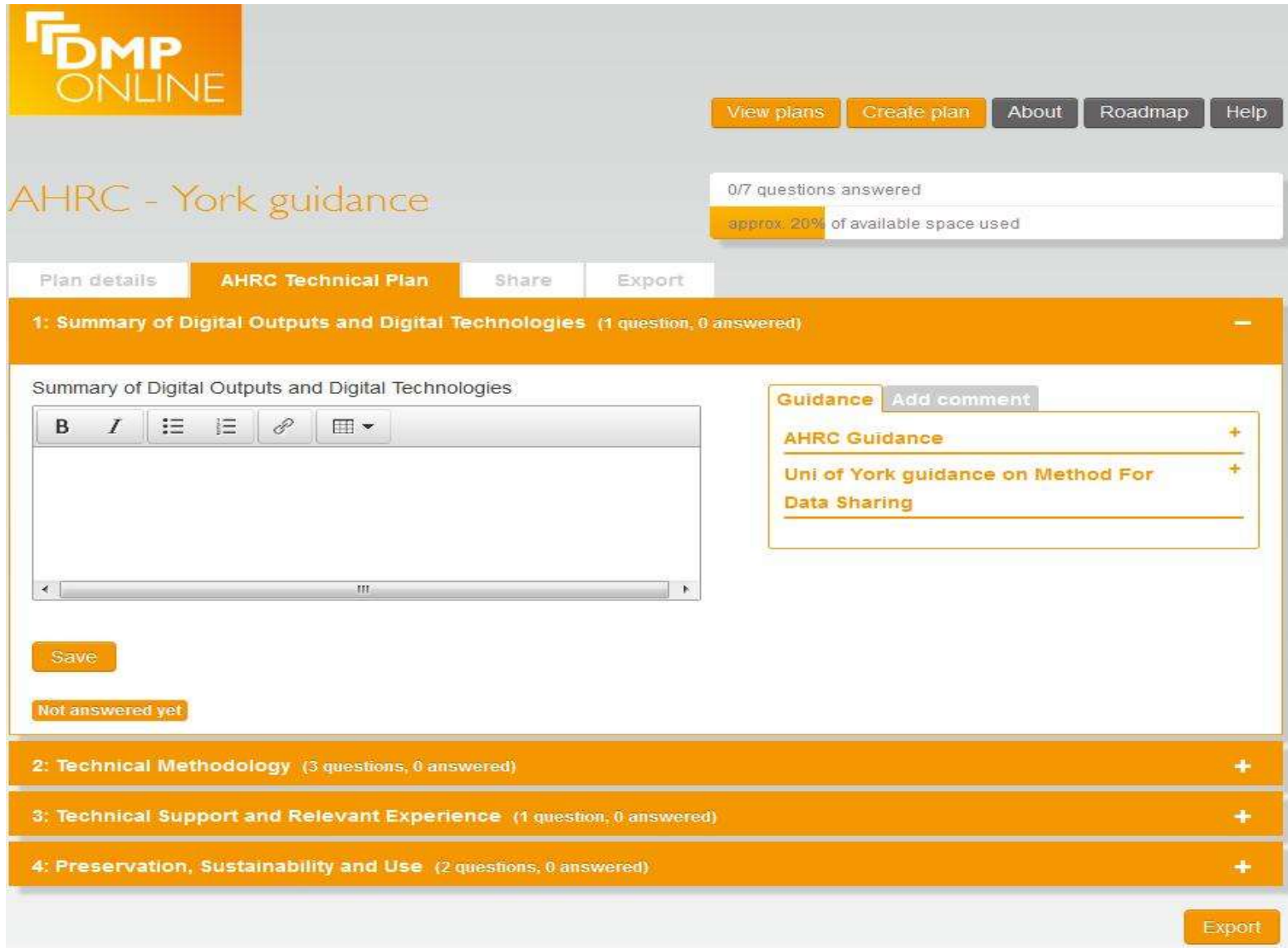
Remember me

Sign in

[Or, sign in with your institutional credentials](#) (UK users only) 

Sign up

New to DMPonline? Sign up today.



The screenshot displays the DMPonline web interface. At the top left is the 'DMP ONLINE' logo. To the right are navigation buttons: 'View plans', 'Create plan', 'About', 'Roadmap', and 'Help'. Below the logo, the title 'AHRC - York guidance' is shown. A progress indicator indicates '0/7 questions answered' and 'approx. 20% of available space used'. A tabbed interface shows 'Plan details', 'AHRC Technical Plan', 'Share', and 'Export'. The main content area is titled '1: Summary of Digital Outputs and Digital Technologies (1 question, 0 answered)'. It features a rich text editor with a toolbar containing bold, italic, list, link, and table icons. Below the editor is a 'Save' button and a 'Not answered yet' status. To the right is a 'Guidance' section with an 'Add comment' button and a list of links: 'AHRC Guidance', 'Uni of York guidance on Method For Data Sharing', and 'Data Sharing'. At the bottom, a list of sections is visible: '2: Technical Methodology (3 questions, 0 answered)', '3: Technical Support and Relevant Experience (1 question, 0 answered)', and '4: Preservation, Sustainability and Use (2 questions, 0 answered)'. An 'Export' button is located at the bottom right.

Further help and information



Help at your institution

Leeds

Website: <http://researchdata.leeds.ac.uk>

Enquiries: researchdataenquiries@leeds.ac.uk

Sheffield

Website: www.sheffield.ac.uk/library/rdm

Enquiries: rdm@sheffield.ac.uk

York

Website: www.york.ac.uk/rdm

Enquiries: lib-research-support@york.ac.uk

Training

MANTRA: Research data management training
(University of Edinburgh)

<http://datalib.edina.ac.uk/mantra>

Leeds: Research@Library

<http://library.leeds.ac.uk/researcher-training>

Sheffield: Doctoral Development Programme

www.sheffield.ac.uk/ris/pgr/ddpportal

York: Researcher Development

www.york.ac.uk/admin/hr/researcher-development

Summary



A final word from you?

Questions

Comments

- things you've learnt
- what you need to investigate further
- what's been useful
- what hasn't ...

Summary

Publicly funded research should be for the common good

Managing research data is part of good research practice

- The data allows you to justify your research findings
- Enables you to more easily find and reuse your research data
- Managed data can be shared with others

A research data management plan helps in achieving this

Be clear about who is responsible for research data and its management

We need your

FEEDBACK

please fill out the
evaluation form provided

<http://bit.ly/wrocah-22Jun15-FB>



References

- Digital Curation Centre. *Digital Curation Centre web site*. [Online]. [Accessed 17 Feb 2015]. Available from: www.dcc.ac.uk
- UK Data Archive. *UK Data Archive website: create and manage data*. [Online]. [Accessed 17 Feb 2015]. Available from: www.data-archive.ac.uk/create-manage.

Data citation references

- Bertil Dorch. 2012. On the Citation Advantage of linking to data. <https://halshs.archives-ouvertes.fr/hprints-00714715/>
- Piwowar HA, Vision TJ. 2013. Data reuse and the open data citation advantage. PeerJ 1:e175 <http://dx.doi.org/10.7717/peerj.175>
- Wicherts JM, Bakker M, Molenaar D. 2011. Willingness to Share Research Data Is Related to the Strength of the Evidence and the Quality of Reporting of Statistical Results PLoS ONE 6(11):e26828. <http://doi.org/10.1371/journal.pone.0026828>