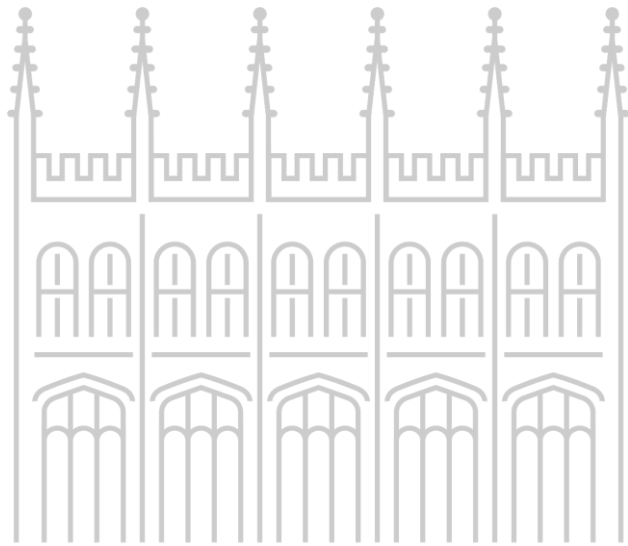


# iSkills: Working with Sensitive or Confidential Research Data in the Sciences & Medical Sciences

Michaelmas Term

November 2022



BODLEIAN  
LIBRARIES



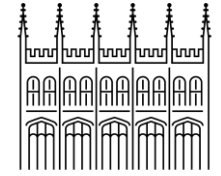
John Southall  
Bodleian Data Librarian  
[john.southall@bodleian.ox.ac.uk](mailto:john.southall@bodleian.ox.ac.uk)



John Pilbeam  
Web/Digital Officer  
[john.pilbeam@sbs.ox.ac.uk](mailto:john.pilbeam@sbs.ox.ac.uk)

# Overview

- Consider sensitive data in relation to;
  - Storage
  - Preservation
  - Legislation
  - Typical research practices



**BODLEIAN**  
LIBRARIES  
UNIVERSITY OF OXFORD

- Consider sensitive data in relation to;
  - Storage
  - Preservation
  - Legislation
  - Typical research practices

# Data Storage



- Securely storing the data is key
- During research when data is active / live
- Requires appropriate secure handling and storage
  - Use approved tools – OneDrive for Business
  - <https://help.it.ox.ac.uk/which-onedrive>
  - Seek advice from department (local solutions) and RDO



## Welcome to the Research Data Oxford website



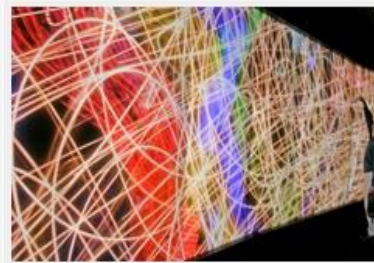
### About RDM

Overview of research data management and funder policies.



### Working with data

Data management day-to-day and at the project planning stage.



### Sharing data

Sharing, licensing, depositing, and citing your data.



### Tools, services, and training



### Research data glossary



### Oxford research data blog

## ORA-Data



Not sure if you're ready?  
See the [Pre-deposit checklist](#)

[University research data policy](#)

[What does my funder expect?](#)

[Data management planning](#)

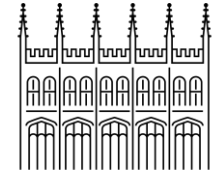
[A to Z site index](#)

[Contact us](#)

Recent blog posts

# Data Storage

- Securely storing the data is key
  - Especially in the short term / during research activity
    - But only **one** part
- How to appropriately manage sensitive data
  - Moving material around
  - Honouring agreements made
  - Avoid convenient but unapproved short term tools –  
Dropbox, email, Amazon Drive etc.



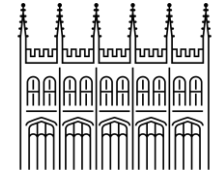
**BODLEIAN**  
LIBRARIES  
UNIVERSITY OF OXFORD

- Consider sensitive data in relation to;
  - Storage
  - Preservation
  - Legislation
  - Typical research practices

# Data Preservation

- Consider efficient access for you
  - In the long term
  - After analysis/ write up
  - If/When OneDrive is no longer available?
- Issues around sensitive data
  - Is preservation covered by consent/ agreements
  - Is documentation created to provide context
  - Are ethical and legal implications being considered
  - Is there evidence ( documentation ) of this

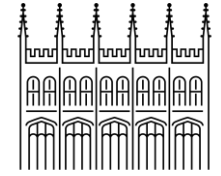




**BODLEIAN**  
LIBRARIES  
UNIVERSITY OF OXFORD

- Consider sensitive data in relation to;
  - Storage
  - Preservation
  - Legislation
  - Typical research practices

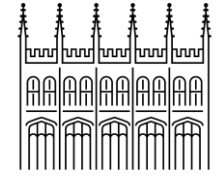
# Legal Regulation - GDPR



BODLEIAN  
LIBRARIES  
UNIVERSITY OF OXFORD

- General Data Protection Regulation (GDPR)
- Addresses handling/processing of personal data
- Information Commissioners Office ICO definition of personal data
- “If it is possible to identify an individual directly from the information you are processing, then that information may be personal data.”
- <https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/key-definitions/>

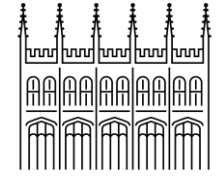
# Legal Regulation - GDPR



BODLEIAN  
LIBRARIES  
UNIVERSITY OF OXFORD

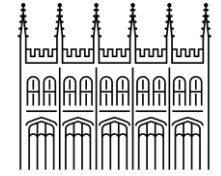
- ... But notice the wording of the ICO definition of personal data
- “*If* it is possible to identify an individual directly from the information you are processing, then that information *may* be personal data.”
- Consider this carefully and be prepared to defend *your* definition.
- Interpretation of the regulation in the context of wider RDM decisions

# GDPR Exemptions



BODLEIAN  
LIBRARIES  
UNIVERSITY OF OXFORD

- Non-commercial / Non-administrative use
  - “Research occupies a privileged position within the Regulation. Organizations that process personal data for research purposes may avoid restrictions on secondary processing and on processing sensitive categories of data (Article 6(4); Recital 50). *As long as they implement appropriate safeguards*”
  - “...these organizations also may override a data subject’s right to object to processing and to seek the erasure of personal data” (Article 89).



**BODLEIAN**  
LIBRARIES  
UNIVERSITY OF OXFORD

- Consider sensitive data in relation to;
  - Storage
  - Preservation
  - Legislation
  - Typical research practices

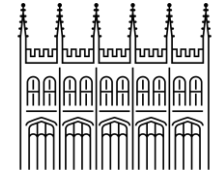
# Three General Approaches



Characteristic of handling personal or sensitive data

- Destroy
- Anonymise
- Restrict

# Data Destruction



BODLEIAN  
LIBRARIES  
UNIVERSITY OF OXFORD

- During or after a project
- Make a good case for this
  - Full or partial destruction?
- Is there an intended retention period for personal data?
  - “The storage limitation principle states that we must not keep data longer than necessary for the *purposes* for which it was collected.”
  - <https://compliance.admin.ox.ac.uk/retention-schedules#collapse1098971>
- But how are purposes being defined?

# Choosing Data Destruction

## Where destruction route chosen

- Must be planned and secure
- Use appropriate/approved tools
  - Eraser - Blancco - Disk Utility (Mac)
  - Infosec / RDO / Departmental recommendation
- But.. This assumes nothing is being preserved
  - Might be a project objective
  - Might be a funder/ publisher obligation
- So there are other strategies and approaches



# Anonymisation



## Removing certain sections or variables

- When? During or after a project?
- Replacement / Pseudonyms – data blurring
- Aggregation – fine grain detail/numbers removed
- **Light** touch is important
  - Specific columns in spreadsheets
  - Limited key identifiers e.g. Names and addresses only

# Choosing Anonymisation

- Perhaps best used for **particular content**
  - Removing postcode columns from spreadsheets
  - Agreed identifiers (names, locations) in transcripts
- But an imperfect solution – too blunt a tool?
- Dangers of data degradation
  - Distortion
  - Debates around rendering data anonymous
  - An actual goal or an aspiration?

# Guidance on Anonymisation



- Literature on anonymisation
  - Techniques of anonymisation and de-anonymisation
  - Ethics of anonymisation
  - Limits of anonymisation
- Online ICPSR guidance on RDM and confidentiality
  - [www.icpsr.umich.edu/web/pages/datamanagement/index.html](http://www.icpsr.umich.edu/web/pages/datamanagement/index.html)
- UK Data Service guidance
  - [ukdataservice.ac.uk/learning-hub/research-data-management/](http://ukdataservice.ac.uk/learning-hub/research-data-management/)

## Oxford Research Archive for Data (ORA-Data): About ORA-Data



Getting Started

About ORA-Data

How to deposit

Deposit checklist

Benefits of deposit

Charges

Open data & funders

### What is research data?



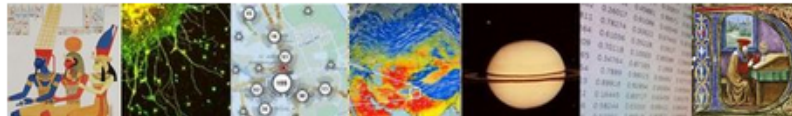
The University of Oxford's [Policy on the management of research data and records](#) defines research data as *'recorded information (regardless of the form or media in which they may exist) necessary to support or validate a research project's observations, findings or outputs'*.

### Oxford research data blog



- [Call for presenters – OxFOS Innovation: Data Circus: Research Data \(and Software\) in Performance](#) [🔗](#) [📅](#) Feb 16, 2022
- [New version of RDM induction](#)

### What is ORA-Data?



ORA-Data is the University of Oxford's digital catalogue and repository for research data, managed by the Bodleian Libraries. It offers a service to archive, preserve and enable the discovery and sharing of data produced by Oxford researchers.

Any type of digital research data – from across all academic disciplines – may be deposited in ORA-Data. We accept any file format.

ORA-Data is aimed at researchers who:

- wish to include an entry for their dataset in the University's catalogue of research data, irrespective of where the data is archived
- need a repository to deposit research data – especially data that underpins publications, and data where the funding body requires archiving and preservation.

ORA-Data sits within the original Oxford Research Archive (ORA) for publications, so that data can be linked easily to – and browsed alongside – related publications.

### Key features of ORA-Data

#### Deposit

### Oxford University Research Archive



**ORA (Oxford University Research Archive)** is the institutional repository for the University of Oxford and is home to the scholarly output of its research members.

Contact us

at: [ora@bodleian.ox.ac.uk](mailto:ora@bodleian.ox.ac.uk), or via our [contact form](#).

### ORA-Data service policies

- [📄 ORA-Data Policy Summary](#)
- [📄 Bodleian Libraries RDM Policy](#)
- [📄 University of Oxford RDM Policy](#)

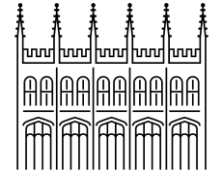
# Restricted Access



## Latest approach builds on growth of data preservation

- Through project websites or dashboards
- Use of archives and repository services e.g. UK Biobank, Health Data Research UK (HDRUK) etc.
- Archiving and sharing data isn't always the same as Open/Public!
- Leave content **but make access harder / restricted**
  - Access is vetted with potential for refusal
  - Project documentation will have clarity on access and usage conditions

# Choosing Restricted Access



BODLEIAN  
LIBRARIES  
UNIVERSITY OF OXFORD

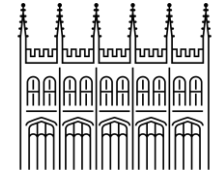
- Best used for **general content** confidentiality?
- Effective or credible policing of restrictions needed
- Requires **planning from the beginning**
- Indicated in consent
- Requires a host archive to act on your behalf

# All three depend on planning



- Plan and **document** the research process
  - Metadata captures decisions with clear requirements
  - How sensitive data will be collected and handled
  - How sensitive data will be managed, preserved or destroyed
- Embrace DMP, CUREC, DPIA and similar as tools to help this

# Aspects of Research Data Management



BODLEIAN  
LIBRARIES  
UNIVERSITY OF OXFORD

- Pilot consent paperwork
  - Does it protect you and participants?
- Think about what could go wrong!
  - Collect unnecessary data
  - Hardware /software failure
  - Security – breaches - theft
- Formal protocols to manage
  - Incidents of data of disclosure / breach
  - Accusations of disclosure



# What next?

- Seek support and advice
- [john.southall@bodleian.ox.ac.uk](mailto:john.southall@bodleian.ox.ac.uk)
- [john.pilbeam@sbs.ox.ac.uk](mailto:john.pilbeam@sbs.ox.ac.uk)

