Digital Data and Archaeology: Management, Preservation and Publishing

Dr Katie Green
katie.green@york.ac.uk
Data Lifecycle

Data Management Plan

Informs all other stages
Data Management Plans
Data management plans should cover the following:

• Description of the data to be collected / created
• Standards / methodologies for data collection and management
• Ethics and Intellectual Property concerns or restrictions
• Plans for data sharing and access
• Strategy for long-term preservation
Data Management

Data Cycles & Management Plans

1. Data Creation
2. Active Use
3. Selection & Evaluation
4. Deposit Data
5. Preservation & Re-Use
Data Cycles & Management Plans

1. Data Creation
2. Active Use
3. Selection & Evaluation
4. Deposit Data
5. Preservation & Re-Use

• What data will I produce?
Data Cycles & Management Plans

1. Data Creation
   - What data will I produce?
     - Text documents
     - Artefact analyses
     - Sample analyses
     - Survey data
     - Drawings
     - Photographs
     - Recorded interviews
     - Etc.

   http://www.jiscdigitalmedia.ac.uk/infokit/file_formats/digital-file-formats

   - Who will own data? Do I need permissions?

2. Active Use

3. Selection & Evaluation

4. Deposit Data

5. Preservation & Re-Use
Data Cycles & Management Plans

1. Data Creation
2. Active Use
3. Selection & Evaluation
4. Deposit Data
5. Preservation & Re-Use

- What data will I produce?
- How will I organise the data?
Data Cycles & Management Plans

1. Data Creation
   - What data will I produce?
   - How will I organise the data?
     - File structure
     - File naming and versioning
     - What file formats will I use?
     - Which software will I use?
     - Roughly how many files?
     - How will I describe and document my data? – METADATA

2. Active Use

3. Selection & Evaluation

4. Deposit Data

5. Preservation & Re-Use
Guides to Good Practice

- Digital Data (general)
- GIS
- CAD
- Geophysics
- AP & Remote Sensing
- Excavation & Fieldwork
- Virtual Reality
- UAV
Data Cycles & Management Plans

1. Data Creation
   - What data will I produce?

2. Active Use
   - How will I organise the data?
     - File structure
     - File naming and versioning
     - What file formats will I use?
     - Which software will I use?
     - Roughly how many files?
     - How will I describe and document my data? – METADATA
     - Do I have to follow any requirements?

3. Selection & Evaluation

4. Deposit Data

5. Preservation & Re-Use
Data Cycles & Management Plans

Guidelines for Depositors
Version 1.3, March 2008

1. Depositing with the ADS
   - 1.1. Why Deposit?
   - 1.2. How to Deposit

2. Creating and Documenting your data
   - 2.1. Part 1: Starting the Project
     ■ 2.1.1. Digital Archive Strategy
     ■ 2.1.2. The need for Metadata / Documentation
     ■ 2.1.3. File Naming Strategy
   - 2.2. Part 2: Creating and Documenting Your Files
     ■ 2.2.1. Overview of Preferred Data Formats
     ■ 2.2.2. Databases and Spreadsheets
     ■ 2.2.3. Geographical Information Systems
     ■ 2.2.4. Geophysics and Remote Sensing
     ■ 2.2.5. CAD and Vector images
     ■ 2.2.6. Raster images
   - 2.3. Part 3: Documenting the Project
     ■ 2.3.1. Creating Metadata Records for Datasets

1. Depositing with the ADS
1.1. Why Deposit?
The Archaeology Data Service (ADS) collects, catalogues, manages, preserves, and encourages re-use of digital resources created by archaeologists. These pages describe the process of deposition and point to useful information about how to do it.

What is in the ADS collection?
The ADS will provide an archival home for any archaeological data of interest to UK archaeologists. The ADS collection’s scope is thus international.

Talk to the digital repository early!
Data Cycles & Management Plans

1. Data Creation
2. Active Use
3. Selection & Evaluation
4. Deposit Data
5. Preservation & Re-Use

• What standards and quality assurance might I use?
Data Cycles & Management Plans

1. Data Creation

2. Active Use

3. Selection & Evaluation

4. Deposit Data

5. Preservation & Re-Use

- What standards and quality assurance might I use?
- How will I share data?
Data Cycles & Management Plans

1. Data Creation
2. Active Use
3. Selection & Evaluation
4. Deposit Data
5. Preservation & Re-Use

• What standards and quality assurance might I use?
• How will I share data?
• How will I backup data?
4. Deposit Data

3. Selection & Evaluation

2. Active Use

1. Data Creation

5. Preservation & Re-Use

Data Cycles & Management Plans

- What standards and quality assurance might I use?
- How will I share data?
- How will I backup data?
- When will I evaluate if my data management is working?
  - Is the file structure / naming understandable to others?
  - Are further data required?
  - Are new data types required?
Data Cycles & Management Plans

1. Create
2. Active Use
3. Selection & Retention
4. Deposit Data
5. Preservation & Re-Use

• What data will I keep?
Selection and Retention strategy
Data Cycles & Management Plans

- What data will be deposited and where?

1. Create
2. Active Use
3. Selection & Evaluation
4. Deposit Data
5. Preservation & Re-Use
Data Cycles & Management Plans

1. Create
2. Active Use
3. Selection & Evaluation
4. Deposit Data
5. Preservation & Re-Use

- What data will be deposited and where?
  - Define the core data set of the project
  - Which data are supplementary?
  - Licences
  - Metadata
  - Where? Trusted Repository!

Talk to the digital repository early!
Data Cycles & Management Plans

Guidelines for Depositors
Version 1.3, March 2008

1. Depositing with the ADS
   1.1. Why Deposit?
   1.2. How to Deposit
2. Creating and Documenting your data
   2.1. Part 1: Starting the Project
      ▪ 2.1.1. Digital Archive Strategy
      ▪ 2.1.2. The need for Metadata / Documentation
      ▪ 2.1.3. File Naming Strategy
   2.2. Part 2: Creating and Documenting Your Files
      ▪ 2.2.1. Overview of Preferred Data Formats
      ▪ 2.2.2. Databases and Spreadsheets
      ▪ 2.2.3. Geographical Information Systems
      ▪ 2.2.4. Geophysics and Remote Sensing
      ▪ 2.2.5. CAD and Vector Images
      ▪ 2.2.6. Raster Images
   2.3. Part 3: Documenting the Project
      ▪ 2.3.1. Creating Metadata Records for Datasets

1. Depositing with the ADS
1.1. Why Deposit?
The Archaeology Data Service (ADS) collects, catalogues, manages, preserves, and encourages re-use of digital resources created by
archaeologists. These pages describe the process of deposition and points to useful information about how to do it.

What is in the ADS collection?
The ADS will provide an archival home for any archaeological data of interest to UK archaeologists. The ADS collections’ scope is thus international.
Data Cycles & Management Plans

1. Create
2. Active Use
3. Selection & Evaluation
4. Deposit Data
5. Preservation & Re-Use

- Who will be interested in re-using the data?
  - Who will be interested in re-using the data?
  - Is there sufficient information to allow easy re-use of the data?

“The single most useful thing you can do to ensure the long-term preservation of your data is to plan for it to be re-used. 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.

Moreover, studies show that re-use of data is the single surest way of maintaining the integrity of data and tracking errors and problems with it. In short, always plan for re-use”

Prof. Julian Richards, Director ADS.
Why bother?

- provides a **practical starting point** to help structure thoughts on your research/project
- **improves efficiency**
- help others understand the **research process** and how it developed
- helps plan for data **reuse** by others, so the full potential of a research can be realised. Its lifecycle doesn’t end here!
Why bother?

- provides a **practical starting point** to help structure thoughts on your research/project
- **improves efficiency**
- help others understand the **research process** and how it developed
- helps plan for data **reuse** by others, so the full potential of a research can be realised. Its lifecycle doesn’t end here!
- **So we don’t end up with examples like this!**
• Cats-in-a-tent photos….and more
• Cats-in-a-tent photos....and more

• Oral history project with no consent forms and therefore no audio files!

Downloads
We regret that we are unable to offer the archived audio files for this project at present due to copyright restrictions.
• Cats-in-a-tent photos....and more
• Oral history project with no consent forms and therefore no audio files!

Downloads
We regret that we are unable to offer the archived audio files for this project at present due to copyright restrictions.

• Project in the red sea with videos of the rock of Gibraltar
• Cats-in-a-tent photos....and more

• Oral history project with no consent forms and therefore no audio files!

Downloads
We regret that we are unable to offer the archived audio files for this project at present due to copyright restrictions.

• Project in the red sea with videos of the rock of Gibraltar

• Recently an archive with the wrong archaeological site name in all the metadata
Odd one out – which image was included in a deposit but is not on the ADS web site?
Why bother?

• provides a **practical starting point** to help structure thoughts on your research/project
• **improves efficiency**
• help others understand the **research process** and how it developed
• helps plan for data **reuse** by others, so the full potential of a research can be realised. Its lifecycle doesn’t end here!
• shows we take **research integrity** seriously and therefore increases trust in the archaeological community
• **it is good practice**
• **funding bodies** require it!
Creating a DMP: Some useful sources

General guidance on data management and the creation of plans:
http://www.dcc.ac.uk/resources/data-management-plans

DataTrain

Information on data management plans, including blank forms, example, case-studies, learning materials specific to archaeology:
http://archaeologydataservice.ac.uk/learning/DataTrainDownload