DataTrain Archaeology Module 7

Data Management Plans for Post-Graduate Research Projects

Presentation Notes

Written by Lindsay Lloyd-Smith

(July 2011)

Slide 1 Data Management Plans for Post-Graduate Research Projects

This brings us on to looking in detail at data management plans.

Slide 2 Data Management Plans for Post-Graduate Research Projects

We will bring together all the issues covered in the previous sessions including:

- Creating and managing digital research data in archaeology;
- The idea of a Data Lifecycle;
- Practical issues about file structure, naming, and formats,
- The issues of rights and research data;
- E-theses and supplementary data; and lastly,
- How we go about archiving digital data,

And, very quickly run through each of the earlier sessions and re-visit the exercises.

The idea is that you can use the material from these earlier exercises as a base to write a full data management plan for your post-graduate research project.

There is a template form for a Data Management Plan which you can use.

Writing a full plan may take a few hours, but it will be a very useful few hours spent and save you many hours or even days of wasted time in the future.

Slide 3 Integrating Data Cycles and Management Plans - 1

A Data Lifecycle model provides both a conceptual and practical framework for planning and carrying out a research project.

There are simple but important questions that need to be addressed at each stage of the cycle, the first being:

- What types of digital data will be produced?

Slide 4 Defining Research Data - Recap

As a gentle introduction to your own Research Data Management you completed this general form. The purpose of this exercise was to begin to think about:

- What the core data set of your research is;
- What will happen to these data in the future;
- Will you have the authority to archive the data; and,
- What ethical or IPR issues might there be.

Slide 5 Integrating Data Cycles and Management Plans - 2

Next we considered the practical day-to-day issues of working with digital data, including file structure, naming, and formats.

Slide 6 File Structure and Naming - Recap

And you started to describe these for your own data.

The purpose of writing a detailed File Structure and Naming Form is consciously to formulate a plan for how the data will be organised.

The idea being that by writing it down, people are more likely to stick to it and be consistent. A common problem many researchers have is maintaining an undefined system simply by trial-and-error.

The process of describing the way you organise your data allows:

- You to understand the structure of your own data;
- Allows others to understand your data;
- Establishes good working habits.

Re-appraisal of how you look after your data is key and should be addressed after the first year. It is recommended at the beginning of the second year of a project, you go through your management plan, evaluate what is working well, and update it with amendments.

Selecting what data we keep and what we chuck comes down very much to an issue of version control – and whether we keep all the digital data that we are given by others during the course of our research.

Research – particularly doctoral research projects – change. We hardly ever actually answer the questions we set out in our first proposals and very often we end up analysing completely different data as well. Also, very often we end up collecting all sorts of extra data that is not included in the thesis, and this raises the question of what we do with all this stuff. Indeed, what are we allowed to do with it?

Slide 7 Integrating Data Cycles and Management Plans - 3

Moving on from the practicalities, a complete Data Management Plan encompasses all of the stages around a Data Lifecycle. It is formal project document that:

- Formalises the definition of your research data;
- Documents the contextual and technical details of your data;
- Explains your file structure and naming protocols;
- Outlines your proposal for what will happen with your data in the future;
- Describes what plans you have for sharing your data.

It is a good to have an idea at the start of your project of what the key deliverables are, in terms of what goes into your PhD, what data will be included as appendices and in what format. Here you have to consider whether there are any legal or ethical issues regarding the digital data.

- Are you allowed to make the data open access to others?
- What are private data and what are public data?

A question to ask concerning long-term preservation and potential re-use of data is:

- Who is likely to be interested in the data in the future?

The answer may affect how you format and present the data for possible re-use.

Slide 8 DMP Template Form

Various Data Management Plans are available online. The key elements from these that are applicable for Post-Graduate Projects in archaeology have been taken and reduced into two simplified forms. One of which, the File Structure and Naming form, you have already written a first draft of.

The main Data Management Plan deals with the whole research project in general terms. The form is organised so that it follows the data lifecycle model.

First you define what data will be studied, and how the data will be documented. You should describe the timetable for the data management tasks over the three years of the project.

There is then a check box reminder to make sure that the details on how the data will be organised have been completed.

The last three sections put in place plans for what will happen to the data at the end of the project. In particular they make clear arrangements for any ethical or legal issues relating to the data and to think about long term preservation and potential re-use of the data.

The more detail you can put down the better, but at the same time try to keep things simple. What may seem logical and self-explanatory to you is often utterly confusing and bizarre to others. A key point to bear in mind is that somebody else, who knows nothing about your work, should be able to read it and be able to navigate around, understand and re-use the data in the future.

Lastly, it is important to recognise that we will change the way we organise our data as our research progresses. It is recommended that Data Management Plans are reviewed every year and updated.

Presentation suggestion
Along with the template form there is a prompt sheet which can be given to the students at the same time.

The Data Management Plan and File Structure and Naming form can be downloaded and completed digitally.

Slide 9 Back to the Future

Lastly it is worthwhile to end on this slide again. As the ADS says:

"*The best way to help long-term preservation of your data is to plan for its re-use"*.

To be able to do this we need to have the authority to deposit and share the data. It is essential, at the beginning of a project, to understand the issues surrounding the ownership of, and permissions to use, deposit and share data.

Slide 10 Acknowledgements

Module 7 Data Management Plans for Post-Graduate Research Projects

Written by Lindsay Lloyd-Smith (2011)

Acknowledgments
This material was created by the JISC-funded DataTrain Project based at the Cambridge University Library.

Project Manager: Elin Stangeland (Cambridge University Library)
Project advisors: Stuart Jeffrey (Archaeology Data Service), Sian Lazar (Department of Anthropology, Cambridge University), Irene Peano (DataTrain Project Officer: Social Anthropology), Cameron Petrie (Department of Archaeology, Cambridge University), Grant Young (Cambridge University Library), and Anna Collins (DSpace@Cambridge Research Data and Digital Curation Officer).