

Data management plans: what are they and how libraries can help.

Anton Angelo, Peter Lund, University of Canterbury Library



Introduction

Formal data management plans (DMPs) are becoming required for projects funded through UK and EU sources. Data sets are increasingly being expected to accompany papers in order to encourage re-analysis and scholarly transparency, and DMPs outline expectations of how datasets will be presented as final research outputs. Considerations for a DMP include : format, privacy, licensing and re-use restrictions, archiving, persistent identification and compliance to funder's and institutional policies. Support for researchers creating DMPs can come from collaboration between University librarians, research support offices and ICT departments to create services that fit research needs. Some Universities (Duke University, M.I.T., Edinburgh, Oxford) now have data librarians to support this work. At University of Canterbury we are investigating how UC Library can support researchers in development of DMPs. Up-skilling and advocacy will be required in order for Libraries to provide Library Guides, training and advice. Promotion in the research community will be crucial.

What is a data management plan?

"A Data Management Plan is a project document which describes the data (or similar evidence) that a project will collect, how it will be stored during the project, how it will be archived at the end of the project and how access will be granted to it where appropriate."

<http://www.bath.ac.uk/research/data/planning>

Backup and Security

During the research process data gathered needs to be appropriately stored and secured. Experimental design might dictate that researchers may not have full access to all the data (e.g. anonymous randomised control trials). Off site automated backups need to be just as secure as the original data.

Data Types, Formats and Standards

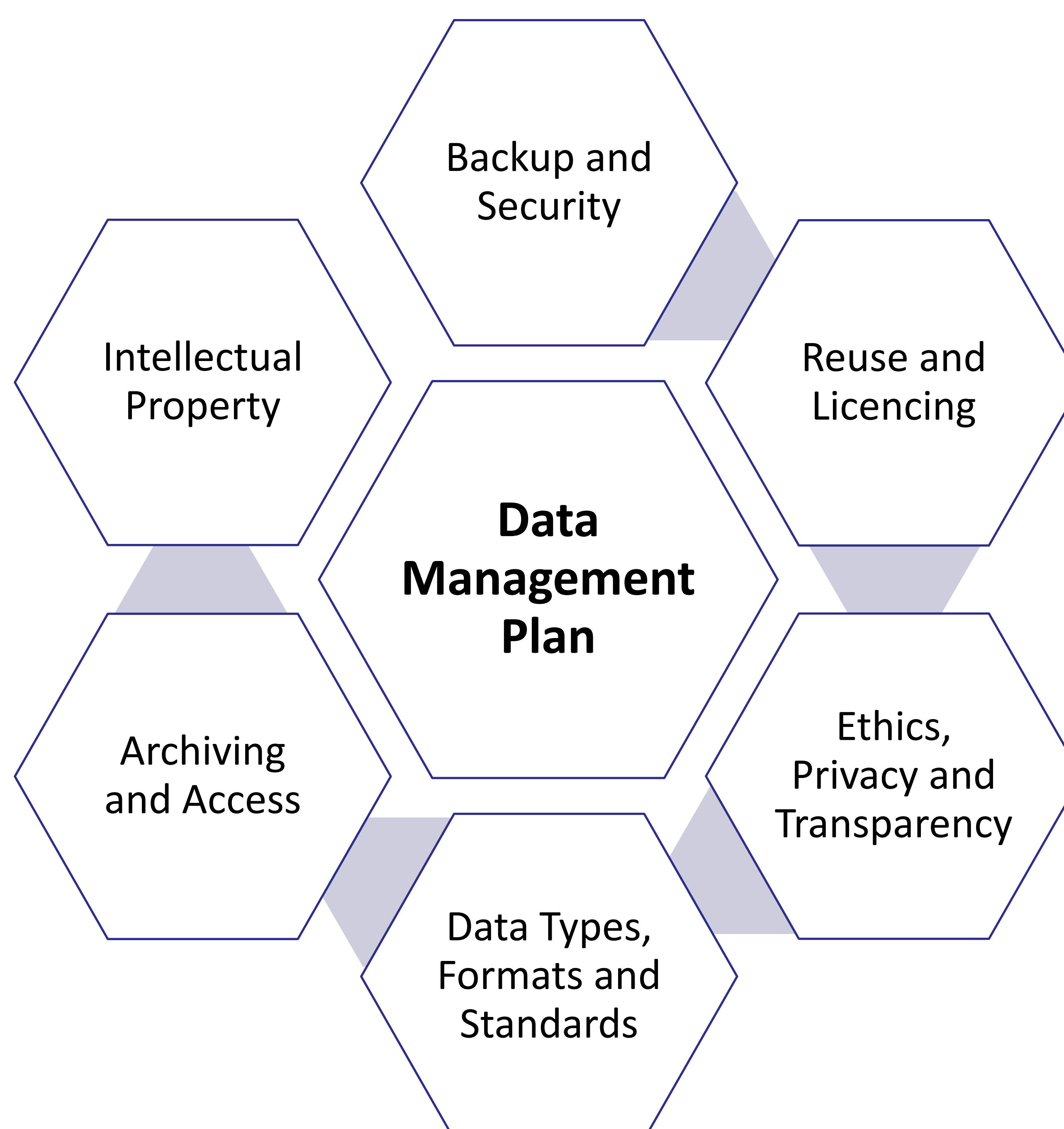
Research data may take the form of audio, video, GIS data or any other format imaginable. A DMP can specify what format data is kept in so to improve collaboration during the research project, as well as ensure that it is accessible afterwards. Simple things like file naming standards can radically improve the findability and usefulness of datasets.

Intellectual Property

Clear ownership of data needs to be established. Whilst expectations are for data resulting from publically funded research to be sharable, embargo periods may be required to protect intellectual property rights. Such embargos or other restrictions will need strong justification. Creating a citation for your data can help.

Archiving and Access

Having an set place in which to store data at the end of a project means that others can find it and analyse it in the future. Techniques such as Digital Object Identifiers or handles enable other researchers to find and attribute your contribution through citations to your data. A data repository will assign an identifier to a dataset, or hold it as part of a collection with the published research output.



Reuse and Licensing

Making your data available to others through an appropriate licence encourages transparency and allowing reuse enables researchers to further your work. Creative Commons has examples of licences that enable sharing and attribution, with specific restrictions commercial use or re-mixing if required.

<http://creativecommons.org/>

Ethics, Privacy and Transparency

Being able to transparently reproduce experimental results is an important part of the scientific method. Reanalysing results from original data provides a similar level of transparency and reliability for data that may not be able to be recreated. Appropriate privacy for sensitive data (e.g. medical or ethnographic) needs to be carefully maintained.

Where can I get help?

Libraries and other institutions provide advice, guides and templates to help with establishing a data management plan. Some good examples are:

- Bath <http://www.bath.ac.uk/research/data> (Full Plans)
- QUT <http://www.tils.qut.edu.au/initiatives/researchsupport/datamange/> (Advice)
- MIT <http://libraries.mit.edu/guides/subjects/data-management/> (Library Guide)
- Research Data Australia <http://researchdata.ands.org.au/> (national data storage)
- Digital Curation Centre <http://www.dcc.ac.uk/resources/data-management-plans> (best practice for standards and formats)



This work is licensed under a Creative Commons Attribution 3.0 New Zealand License.