UNIVERSITY OF PRETORIA

Office of the Vice-Principal: Research and Postgraduate Education

**Project Research Data Management Plan Template**

Document type: Template

Policy Category: Academic

# **PURPOSE**

A Project Data Management Plan documents how researchers plan to collect, store, secure and share their research data. Creating a plan at the beginning of the project will identify and address the main considerations. This will make it easier to identify what the key points are to address requirements of the research funding and publishing bodies.

A good data management plan is essential for successful research, as managing the data effectively across the data lifecycle is necessary for the success of the research project or postgraduate study. In fact, a data management plan is a living document which can be updated as the project develops and the data management strategy is refined.

The following template should be used to develop a Data Management Plan (DMP) to accompany a research proposal. The notes (in italics) provide further context and guidance for its completion. Where substantial data is generated from the research, a detailed DMP will be required, while low impact studies generating small amounts of data may require less detail. Nevertheless, all the topics listed in the template MUST be addressed.

|  |  |
| --- | --- |
| **0. Proposal name** | |
| *Exactly as in the proposal that the DMP accompanies* | |
| **1. Description of the data** | |
| **1.1 Type of study**  *Several lines of text that summarise the type of study (or studies) for which the data are being collected.*  **1.2 Types of data**  *Types of research data to be managed in the following terms: quantitative, qualitative; generated from surveys, clinical measurements, interviews, medical records, electronic health records, administrative records, genotypic data, images, tissue samples,...*  **1.3 Format and scale of the data**  *File formats, software used, number of records, databases, sweeps, repetitions,… (in terms that are meaningful in your field of research). Do formats and software enable sharing and long-term validity of data?* | |
| **2. Data collection / generation** | |
| *The researcher should explain why new data collection and/or long term management is needed.*  *Focus in this template should be on the good practice and standards necessary for ensuring that new data are of high quality and that methods of data processing are well documented.*  **2.1 Methodologies for data collection / generation**  *How the data will be collected and/or generated, and which community data standards (if any) will be used at this stage*.  **2.2 Data quality and standards**  *How consistency and quality of data collection / generation will be controlled and documented, through processes of calibration, repeat samples or measurements, standardised data capture or recording, data entry validation, peer review of data or representation with controlled vocabularies.* | |
| **3. Data management, documentation and curation** | |
| *This section should be concise and accessible to readers who are not data-management experts, with a focus on principles, systems and required standards, as follows:*  **3.1 Managing, storing and curating data.**  *Briefly describe how data will be stored, backed-up, managed and curated in the short to medium term. Specify any community agreed or other formal data standards used (with URL references). [Enter data* security *standards in Section 4].*  **3.2 Metadata standards and data documentation**  *Describe what metadata is necessary, regarding the data generated from the research. For example, descriptions of metadata should enable research data to be used by others outside of the primary research team. This may include documenting the methods used to generate the data, analytical and procedural information, capturing instrument metadata alongside data, documenting provenance of data and their coding, detailed descriptions for variables, records, etc.*  **3.3 Data preservation strategy and standards**  *Plans and place for long-term storage, preservation and planned retention period for the research data. Formal preservation standards, if any. Indicate which data may not be retained (if any).* | |
| **4. Data security and confidentiality of potentially disclosive information** | |
| *This section MUST be completed if the research data includes personal data relating to human participants in research.*  *For other research, the safeguarding and security of data should also be considered. Information provided should be in line with the ethical review.*  **4.1 Formal information/data security standards**  *Identify formal information standards with which the study is or will be compliant. An example is ISO 27001. If the organisation is ISO compliant, the registration number should be stated.*  **4.2 Main risks to data security**  *All personal data has an element of risk. Summarise the main risks to the confidentiality and security of information related to human participants, the level of risk and how these risks will be managed. Cover the main processes or facilities for storage and processing of personal data, data access, with controls put in place and any auditing of user compliance with consent and security conditions.*  *It is not sufficient to write not applicable under this heading.* | |
| **5. Data sharing and access** | |
| *Identify any data repository (-ies) that are, or will be, entrusted with storing, curating and/or sharing data from your study, where they exist for particular disciplinary domains or data types. Information on repositories to be sourced.*  **5.1 Suitability for sharing**  *Is the data to be collected (or existing data proposed for use), in the study, suitable for sharing? If yes, briefly state why it is suitable. If No, indicate why the data will not be suitable for sharing and proceed to Section 6.*  **5.2 Discovery by potential users of the research data**  *Indicate how potential new users (outside of your organisation) can find out about your data and identify whether it could be suitable for their research purposes, e.g. through summary information (metadata) being readily available on the study website, in the UP gateway for population and patient research data, or in other databases or catalogues. How widely accessible is this repository?*  *Indicate whether your policy or approach to data sharing is (or will be) published on your study website (or by other means).*  **5.3 Governance of access**  *Identify who makes or will make the decision on whether to supply research data to a potential new user.*  *Indicate whether the research data will be deposited in and available from an identified community database, repository, archive or other infrastructure established to curate and share data.*  **5.4 The study team’s exclusive use of the data**  *UP’s requirement is for timely data sharing, with the understanding that a limited, defined period of exclusive use of data for primary research is reasonable, according to the nature and value of the data, and that this restriction on sharing should be based on simple, clear principles*. *What are the timescale/dependencies for when data will be accessible to others outside of the research team? Summarize the principles of the current/intended policy.*  **5.5 Restrictions or delays to sharing, with planned actions to limit such restrictions**  *Restriction to data sharing may be due to participant confidentiality, consent agreements or (Intellectual property) IPR. Strategies to limit restrictions may include data being anonymised or aggregated; gaining participant consent for data sharing; gaining copyright permissions. For prospective studies, consent procedures should include provision for data sharing to maximise the value of the data for wider research use, while providing adequate safeguards for participants. As part of the consent process, proposed procedures for data sharing should be set out clearly and current and potential future risks associated with this explained to research participants.*  **5.6 Regulation of responsibilities of users**  *Indicate whether external users are (or will be) bound by data sharing agreements, setting out their main responsibilities.* | |
| **6. Responsibilities** | |
| *Apart from the PI, who is responsible at your organisation/within the consortium for:*   * *study-wide data management* * *metadata creation,* * *data security* * *quality assurance of data.* | |
| **7. Relevant institutional, departmental or study policies on data sharing and data security** | |
| *Please complete, where such policies are (i) relevant to your study, and (ii) are in the public domain, e.g. accessible through the internet.*  *Add any others that are relevant* | |
| **Policy** | **URL or Reference** |
| Data Management Policy & Procedures |  |
| Data Security Policy |  |
| Data Sharing Policy |  |
| Institutional Information Policy |  |
| Other: |  |
| Other |  |
| **8. Author of this Data Management Plan (Name)** and, if different to that of the Principal Investigator, their **telephone & email contact details** | |
|  | |