

First-Year Review Data Management Plan Assessment Checklist

This document is a guide for assessing first-year review data management plans (DMPs). Not all questions will be applicable to all students since each DMP is specific to its own project. Therefore when considering this checklist it is recommended to answer each question with yes/no/not applicable.

1. Types of data, formats, and expected volume

- Has the student described the new data/software they will create? Details can include:
 - Quantitative/qualitative
 - Type of generation (from survey, clinical measurements, interviews, recordings, images, samples, focus groups, demographics, etc.).
- Does the plan specify whether any data include sensitive confidential or personal information?
- Does the student specify details for existing secondary data/sources used?

2. Collection and transfer

- Has the student specified the method of collection/generation for each type of data?
 - Has the student described controls and documentation for managing the method? For example,
 - Calibration
 - Repeat samples or measurements
 - Standardised data capture or recording
 - Data entry validation
 - Peer review
 - Representation with controlled vocabularies, etc.
- Does the student specify when personal data can be anonymised/pseudo-anonymised? What is the earliest possible stage for anonymisation?
- Is there a plan for any transfer of data away from the point of collection?
 - The student may want to contact dataprot@st-andrews.ac.uk for advice if data will transfer outside of the European Economic Area (EEA)?

3. Storage, backup, and access

- Does the student specify how and where the data will be stored?
 - If non-University storage is used, does the student provide a justification?
 - If the data is sensitive, does the student consider appropriate storage and access security measures? For more information, see the [Research data protection guidance page](#).
- Does the student describe how they intend to back-up their data?

4. Documentation

- Does the student plan to use documentation to make the data useable by others both in and outside the team? For example,
 - Documenting the methods used to generate the data
 - Analytical and procedural information
 - Capturing instrument metadata alongside data
 - Documenting provenance of data and its coding
 - Detailed descriptions for variables, records, etc.

5. Main risks to data security and how these will be mitigated

- Does the student summarise the main risks to confidentiality and security of information?
- Does the student plan for managing these risks?
- Does the student discuss storage and security measures (might have been covered in section 3)?
 - Does the student consider who can access commercial, confidential, and personal data?

6. Long-term data retention, preservation strategy, and destruction

- Does the student plan for long-term storage, preservation, and retention of research data?
- If any data should not be retained, does the student indicate this and provide a justification?

7. Sharing and publication strategy (including restrictions and delays)

- Does the student describe if, where, and in what form the data will be made publicly available?
 - Does the student consider whether all or part of the data should be shared, and why?
 - Where making data publicly available is not possible, does the student provide a justification?
- Does the student consider stakeholder policies in their data sharing approach? This could include institutional, funder, and publisher policies.