​​**Data Management Checklist**

**How to use this checklist:**

This checklist is intended to help you get started integrating data management into your research practice. This checklist can help you identify gaps and communicate elements of data management to members of your research team. It is recommended that you apply this checklist to an individual research project as practices and procedures may vary considerably between projects.

The focus of this checklist is on practices implemented by you and your collaborators. When managing research data, it is extremely valuable to seek out and follow community standards and practices rather than developing practices from scratch. For assistance identifying standards or practices that may have been developed for your research area, contact [organization-specific or community-specific data management expertise].

This checklist is not intended to cover every single aspect of research data management. For example, the current iteration does not cover practices or strategies related to quality control in depth. It is likely that certain data management practices that are specific to your research, your type(s) of data, and your needs as a researcher are not covered. It is also possible that certain items on the checklist will not apply to the specific type(s) of data you are working with.

Research data management exists along a continuum and following best practices often involves adopting new guidance, procedures, and techniques as they are developed. So it is worth revisiting this checklist with project team members periodically.

Please feel free to modify this checklist or adapt it to better fit your needs.

If you have any questions about this checklist or research data management more generally, contact [organization-specific or community-specific data management expertise].

**Project Name:**

**Project Team Members:**

**Last Date Reviewed:**

If you and your research team can confidently check off the following three prompts for your project, you are probably doing a reasonable job of managing your data. Because these prompts are so broad, we would encourage you to discuss *how* you are implementing related practices and strategies and working through the rest of the checklist with your team.

* Every member of the research team is able to find and use the data, code, documentation and other materials related to this project.
* Another researcher who works in the same field would be able to find and use the data, code, documentation and other materials related to this project.
* We believe we will be able to find and use the data, code, documentation, and other materials related to this project ten years from now.

The following items relate to specific data-related practices. If you and your research team can confidently check off the following prompts, you have started the process of integrating good data management practices into your project. In this context, standardized practices are those that are consistently employed by every member of the research team.

*We have done the following at the beginning of our project:*

* We have reviewed all applicable policies related to our data (including from our institution, funder, potential publisher, etc)
* We have read through and understand other relevant agreements, licenses, or other requirements related to our data (e.g. data use agreements, IRB or funder policies).
* We have sought out community standards and best practices related to our data.
* We have discussed the intended products of this project (papers, datasets, software tools, etc) and have decided to what extent we will be making our data and other materials available to others.

*We have a plan:*

* We maintain documentation that describes the type(s) of data we are collecting/analyzing/working with over the course of the project as well as details about materials that are needed to understand and use the data (documentation, code, etc).
* We maintain documentation that describes the specific data management practices (e.g. file naming, formats, and standards, backing up data) employed throughout the course of this project.
* We maintain documentation that outlines the roles and responsibilities of individual team members related to managing data (e.g. maintaining good documentation, following file naming conventions, etc) as well as who is ultimately responsible for ensuring the data is properly managed throughout the course of the project and following its conclusion.
* Members of the research team have access to the above documentation and review it periodically.

*We are keeping our data organized:*

* We have a standardized set of practices related to saving datasets and other project materials while we are working with them (e.g. digital data is saved on a lab server).
* Our practices related to saving data are in line with [our organization’s] risk classification system and, when possible and appropriate, include multiple backups.
* We have standardized conventions for naming project-related objects and files (including datasets) that enable us to quickly identify the materials we are looking for.
* We have standardized systems for organizing project-related objects or files that enable us to easily find the materials we are looking for (e.g. a standardized file structure).
* When applicable, we have standardized systems for naming and organizing information within our data files (e.g. standardized variable names, tidy spreadsheets).
* Our practices related to saving, organizing, and describing data files have been optimized to facilitate quality control.
* Our practices related to saving, organizing, and describing data files are in line with community standards and best practices.

*We are keeping good records:*

* We maintain documentation that describes how we keep datasets and other materials organized while we are working with them (e.g. naming conventions, file structures, etc).
* We have standardized procedures for documenting the structure and contents of individual data files (e.g. maintaining codebooks, data dictionaries, etc).
* We have standardized procedures for documenting project-related decisions, steps, procedures, and workflows (e.g. maintaining protocols, lab notebooks, etc).
* We have standardized procedures for saving and versioning research-related code and other elements of the research process (e.g. workflows, software containers).

*We have done (or will do) the following before the end of the project:*

* When necessary and appropriate, datasets and related materials are converted into a form suitable for long-term storage or archiving (e.g. open file formats for digital files).
* If we have decided to share our data, we have uploaded it to a suitable data repository to make it discoverable and FAIR together with any documentation and materials that are necessary to make use of it.
* We have moved project-related data, documentation, and other materials to a location suitable for long-term storage or archiving that we are able to access when necessary.

*We are checking up on ourselves:*

* Members of the research team are actually following the practices and procedures we have decided upon.
* Study documentation is updated regularly to reflect any changes to data management-related practices and procedures.
* We have adopted community standards and practices whenever possible,
* We have established procedures for onboarding new team members about our data management practices, educating members about changes to existing practices, and managing data as team members move onto new projects.