



Research Methods in Science in Engineering

Data Management

Slide 1

1

What is Data Management?



Data management describes what data is to be collected, how the data will be managed and stored, and how the data will be shared with others.

Slide 2

2

Data Description

- What is the purpose of the research?
- What data will be collected and in what format?
- How much data will be collected?
- Will the data change over time?
- Are multiple people or organizations collecting data?
- Who is responsible for data management of your research?



Organization & Storage

- How will you document and organize the data?
- Is your data classified or protected in any way?
- Does your data entail metadata (date, time, location, etc.)
- What file formats will be used? Are they standard or proprietary?
- What directory file naming convention will you use?
- What are your storage and backup procedures?
- What tools or software are needed to view and use the data?
- How will you prevent accidental deletion of data?



File Naming Convention (FNC)

One example of a file naming convention:

[Project]_[Document Title]_[Date]_[Optional Version].[File Extension]

“ResearchMethods_Homework1_30Feb2030_v2.docx”

General Guidelines for FNCs:

- Identify everything important and define a naming convention that includes all of it.
- Be consistent and rigorous.
- Consider doing the same for folders and files.
- Use “_” instead of “ ” because some archival tools cannot handle spaces.
- Never use words like “final” in the file name.

Storage & Backups

- Backup all of your data at least weekly.
- Backups should be stored offsite and be off-network to protect against cyber attacks.
- Limit access and permissions to the backups to prevent accidental deletion or overwriting.
- Include redundancy in your backup system (RAID, etc.)

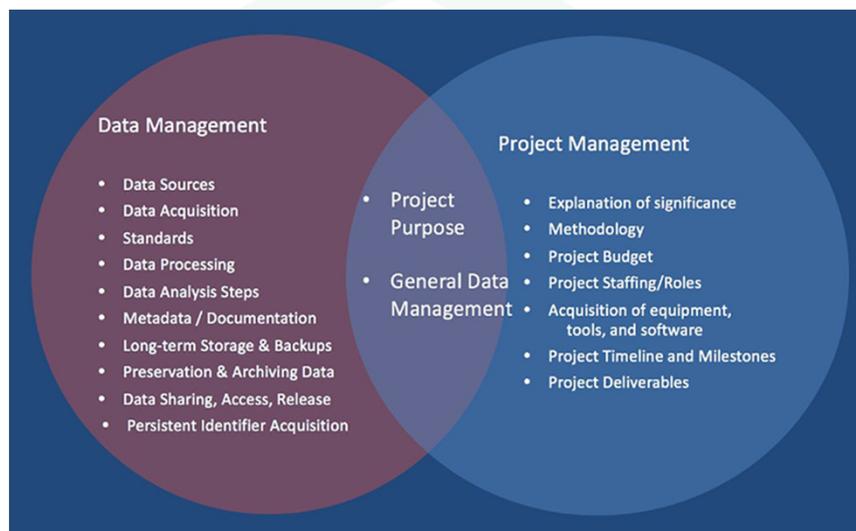


Access to the Data

- What data will be shared?
- When will the data be shared?
 - Is IP being developed, publications being written, etc.?
- How will the data be shared?
- Who will the data be shared with?
 - Any patents or licensing restrictions? Embargos?
- Are there any privacy, ethical or confidentiality concerns?
- Who holds the intellectual property rights to the data?
- Is the data open-source? Is re-use by others permitted?



Data Management Vs. Project Management



<https://www.usgs.gov/media/images/data-management-vs-project-management-venn-diagram>

Data Management Plan



Many government agencies (like NSF) require a Data Management Plan (DMP) to be included in grant applications.

The government wants to protect and have access to what they paid for.