

Group Project

This assignment offers the opportunity to investigate an emerging area of interest in data management relevant to universities, researchers, and academic librarians. It incorporates questions of both technology and policy. It is a domain with few established practices, and no proven best practices. There is a broad range of acceptable answers; the focus is to demonstrate your understanding of data management.

Deadline & Submission

The assignment is due at 11:59pm on April 10. It is to be completed in groups of 4 members (you may build your own groups) and submitted in one group member's GitHub account (as discussed in class). The group has to send me an email, indicating which GitHub account will have the final submission. The group member with the final submission file will NOT get any extra points for this. The final submission should be in PDF format.

Instructions

As discussed in class, major funding agencies are asking or requiring that researchers receiving public funds preserve and share the products of primary data collection. This data includes "...quantitative social, political, and economic data sets; qualitative information in digital format; experimental research data; still and moving image and sound data bases; and other digital objects used for analytical purposes."¹ There are many good reasons to carefully manage collected research data: to contribute to public knowledge; to enable attempts to reproduce experiments and analyses; to improve collaboration; for future reference; to prove originality if questioned; to make the initial analysis easier and better structured; and many more. While this project approaches this from the perspective of research data, very similar issues exist in the general organization domain.

For this project, your group has been hired as research data consultants at a small university in Nova Scotia. You report to the University Librarian who has hired you to consult with three university researchers on their research data needs for upcoming projects and prepare a data management report/plan for them. Four researchers have signed up to have data management plans made. The Librarian has informed you that you are free to choose three of the four to work with.

Your consultation will occur in two stages. First, the Librarian has provided you with a brief written document answering a standard set of questions about each of the researchers in question. Second, you may ask follow-up questions.

For each case, produce a “Data Management Consultation” report. The audience for this report is the researcher, who will use it as the basis for their actual data management plans, and your manager, who will use it to identify the resources required. There is no standard template, but it should contain the following information:

- What data will be created or collected (type, size, format, etc.)
- What licenses apply to the data
- What facilities and equipment will be required (hard disk space, backup server, central repository, off-site repository, etc.)
- What data management practices (backups, storage, access control, archiving etc.) will be used
- Who will own and have access to the data
- Which data will retain value after the life of the project
- What metadata and linked open data strategies will be employed
- How will its reuse be enabled and long-term preservation ensured after the original research is completed
- How much will the storage of this data cost (cloud and/or hard drives)

Helpful hints

There are two important notes: first, the cases express answers like a researcher might, while this plan should use the language of a data/information management professional. Second, as an information professional, you have a responsibility to advocate for responsible data management practices, regardless of the preferences of the researcher. Put more directly, if you think a researcher is misguided in terms of their answers, you should try to convince them otherwise.