Apr 9th, 12:00 PM

Developing a Data Management Plan for a Corporate Laboratory: Using a Case Study Method for Teaching

Eric M. Kuzma
*Simmons College*

Follow this and additional works at: [https://escholarship.umassmed.edu/escience_symposium](https://escholarship.umassmed.edu/escience_symposium)

Part of the [Scholarly Communication Commons](https://escholarship.umassmed.edu/scholarly_communication_commons)

This work is licensed under a [Creative Commons Attribution-Noncommercial-Share Alike 3.0 License](https://creativecommons.org/licenses/by-nc-sa/3.0/).

Repository Citation


This material is brought to you by eScholarship@UMMS. It has been accepted for inclusion in University of Massachusetts and New England Area Librarian e-Science Symposium by an authorized administrator of eScholarship@UMMS. For more information, please contact Lisa.Palmer@umassmed.edu.
Abstract
This poster looks at the importance of developing a Data Management Plan (DMP) in the biological sciences. Grant applications increasingly require applicants to attach a DMP to their research proposals to outline the creation, storage, and dissemination of research data and information. One way to understand and assess a corporation’s current DMP is to use the case study method, and this poster will walk through a DMP using this research method. This particular case study looks at the industrial scientific research process and how data is created, named, and stored, with attention to record keeping deficiencies and prospects. It also looks at the issues surrounding animal experimentation.

Why the Case Study Method?
• A case study is a qualitative research method that takes an in-depth look at a particular individual, group, or scenario. The information drawn from the study is then woven into a narrative, representing a realistic setting.
• The scenario will contain an unresolved issue or a conflict. The case study is deliberately ambiguous in order to encourage questions and a lively classroom debate. The instructor should act as a moderator rather than lead the discussion.

Data Life Cycle

How to Conduct Science Interviews Without a Science Background
You do not need to have a deep knowledge of the subject going into the interview. It is a good idea to prepare a basic foundational knowledge of the subject. Do not try and show off all you know – or just learned.

1. Creating a list of open-ended questions is a great way to start a conversation. Open-ended questions give the interviewee the leeway to discuss, in detail, the information they require. You get the information you need without having to learn all the technical lingo to ask the question in the first place.
2. Always ask the interviewee how they would be working if they understood the information they wish to convey. It takes very little effort on the interviewee’s part, you get the information you need, and the interviewee will feel good after the experience.
3. In addition to being nice, relax! Anxiety and nerves are contagious! Do not rush. Silence is good. All parties can use these moments to gather their thoughts.
4. Follow-up questions are OK. Make sure you thoroughly understand the answers as they are explained to you.

Conclusions: Case Study Teaching Points

How to Conduct Science Interviews Without a Science Background

Linking and Finding Data:
The use of lab notebooks continues to be the standard means of recording experiment data points. That information is then transferred to an Excel file and uploaded to the server. The Informatics Lead then reviews the document and uploads it to its respective departmental subfolder according to month of the experiment. There are two issues here. There is no alternative backup to a paper lab notebook. Also, errors can be introduced when copying information from the lab notebook to the Excel file. Using an electronic lab notebook would eliminate that unnecessary step and reduce the chance of a transcription error. This electronic data can be backed up and secured. Test and non-test data should be kept in separate files. Naming conventions should be uniform across all departments.

Complexities Surrounding Industrial Laboratories
The first layer of complexity is that companies are not required to share proprietary data. The issue becomes more complex when funding for a project comes from both grants and corporate sources. These issues can be resolved during the grant application process. For example, the NIH may recognize that there may be circumstances where the other funder requests that public data sharing be restricted. An approved proposal can work out the details between the two funding parties. One can also put in disclosure delays with the timeframe to be agreed upon by all parties.

References

Interview Instrument and Data Management Plan Checklist (DCC, 2013)

A. Administrative Data
1. Name and job title?
2. How many employees and what are their roles?
3. Who is in charge of the grant writing process?

B. Research Process
1. Overview of research?
2. Current experiments?

C. Data Collection
1. What kind of data is being produced and in what formats is it being saved?
2. How is data recorded? Do you keep a lab notebook?

D. Data Organization
1. Do you use a particular naming convention and where do you save your data?
2. In what format is data saved and how it is organized?

E. Data Access and Security
1. What is the data-flow process?
2. With whom do you share the data?
3. Are there any confidentiality concerns?

F. Storage and Backup
1. Where is the data stored and how long is it kept?
2. Are there any storage space concerns?

G. Data Preservation and Posterity
1. How does the lab deal with turnover and who is responsible for their data?
2. Are their concerns with access to proprietary formats in the future?
3. How does your lab ensure long-term preservation of research data?

Data Management Modules and Recommendations (Lamar Soutter Library, n.d.)
Module 1: Overview
Module 2: Types, Formats, and Stages of Data
1. File naming conventions inconsistent
2. Challenges in absence of PI leadership
3. Organization of test and non-test data
3. Lack of synchronization between departmental data resources

Module 3: Contextual Details
1. Non-test data files are not shared
2. Avoid saving data in proprietary formats

Module 4: Data Storage, Backup, and Security
1. Need for on-site data storage
2. Physical lab notebook data backup required

Module 5: Legal and Ethical Issues
1. Ensure that non-animal subjects are treated humanely

Module 6: Data Sharing and Re-Use
1. Information is proprietary for the company and granting agency only

Module 7: Plans for Archiving and Preservation of Data
1. Data should be stored on-site as a backup

DCC Curation Lifecycle Model

Case Narrative
• The study follows the work of a microbiologist working in a corporate laboratory located in Cambridge, Massachusetts. The company works on developing antimicrobial and hydrophilic polymers applied to the surface of endotracheal tubes. This project is jointly funded through the Department of Defense (DOD) and the company.
• These antimicrobial polymers are tested in the lab by exposing them to various concentrations of bacteria. Next steps involve testing on sheep.
• All experimental data is kept in physical lab notebooks and once filed, secured in a locked cabinet onsite.
• There is currently no Principal Investigator.
• At present there are no uniform naming conventions for electronic data files. There is uniformity within but not among departments.
• Files are uploaded to a shared drive, stored in the cloud, and organized by the Informatics Lead located in Michigan.
• The Informatics Lead sorts files according to department and month. Non-test data such as journal articles and white papers are also included.
• Unlike academic labs, results are not published. The information is proprietary and used to develop products and patents. Information is only shared with the granting agency.

Data Life Cycle

Diagram:

How to Conduct Science Interviews Without a Science Background

Linking and Finding Data: