How to Conduct a Data Audit of a Scientific Research Group

By Chris Eaker
University of Tennessee, Knoxville
ceaker@utk.edu
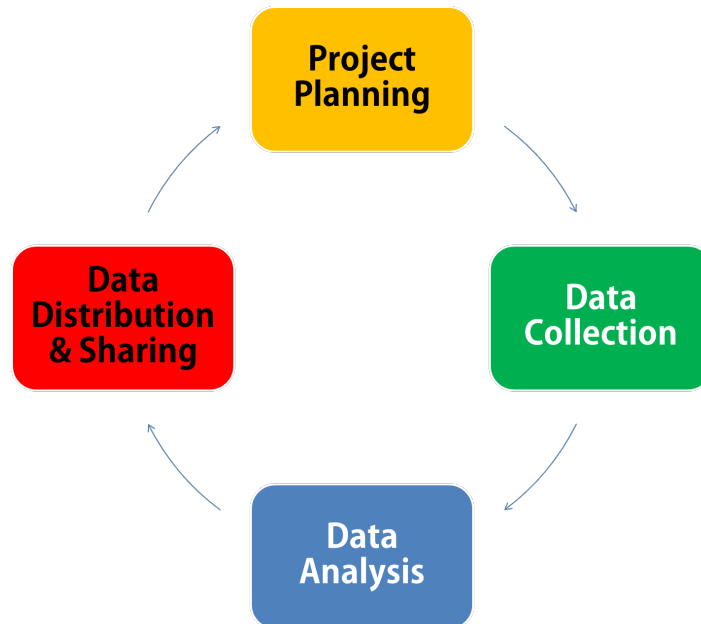
Figure 1. The Data Management Audit Life Cycle

Project Planning

During the project planning phase of the data management audit, the auditor holds introductory meetings with science project managers. The purpose of these meetings is to determine which groups will be audited and be introduced to project personnel who will be participating in the data management audit. The auditor must obtain Internal Review Board approval, if necessary, before conducting any data gathering interviews.

Data Collection

During the data collection phase, the data collected includes, but is not limited to, surveys of project personnel; internal documents, such as reports and project logs; external documents, such as brochures and websites; and demonstrations of tools.

Data Analysis

The data analysis phase of the data management audit includes processing and analyzing the data collected from the previous phase. Tasks within this phase include transcribing interviews, studying each group's workflow, and extracting trends and finding gaps in those workflows. Additionally, the auditor should develop a workflow diagram that visualizes and describes the research process in each group audited. Color-coding each phase of the workflow process on the diagram facilitates easy comparison across research groups. Finally, during data analysis, the auditor must consistently consult with project personnel to determine if the workflow diagrams accurately represent the groups' processes.

Distribution & Sharing

Once the analysis is complete, the auditor must produce a final report that outlines key findings and lays out recommendations for improvements. This report is then shared with the project leaders and other project personnel who will determine how to proceed.

Implementation

Implementation of the recommendations contained within the data management audit report is the goal of the data management audit. The recommendations must be useful, reasonable suggestions to improve efficiency, reduce redundancy, and increase quality. Important areas where recommendations may be made include data management planning, metadata, archiving and preservation, and project tracking.