

International Waters: the case of interdisciplinary research data management

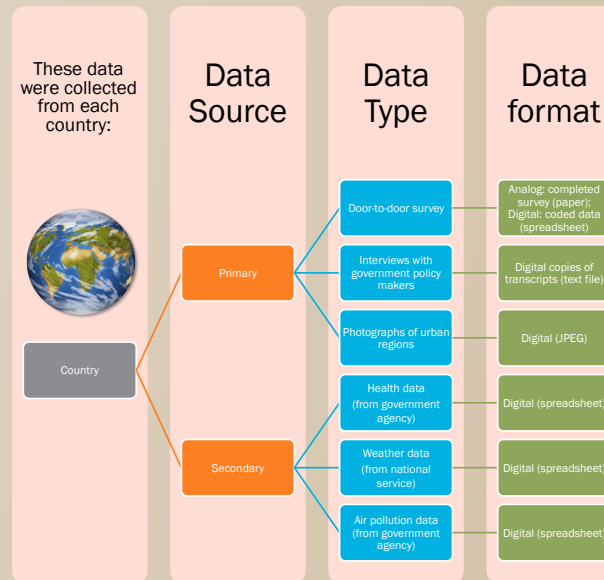
Tiffany Chao
tchao@illinois.edu
University of Illinois,
Urbana-Champaign
GRADUATE SCHOOL OF LIBRARY AND
INFORMATION SCIENCE
The School at Illinois

Introduction

Response to national funding agency policies and mandates to better maintain and curate the data supporting research innovation and progress has been met by the development of data management planning tools and resources. The recommendations and guidance provided by these information resources are not only relevant to funded works but can be adapted more broadly to research projects and associated management practices. Many of these data management services are institutionally supported, however, not all researchers are fortunate enough to have the internal resources or institutional infrastructure for consultation. Given the emergence of publically available guides and tools, this study examines the broad application of these information sources to multidisciplinary research data collected and generated by scientists and the considerations and questions posed in identifying and adopting best practices for data management.

Collaborating across countries

Research situated at the intersections of social science and climate science present innovative projects that span international and disciplinary boundaries. Such areas of focus include human-environment interactions in the context of environmental and global change within urban and metropolitan settings along with relationships between human health and climate processes. Detailed is a preliminary overview of an example project's data, identification of key data management components derived from available resources, and future considerations for international efforts. The collection of data for analysis consists of both analog and digital materials with multi-lingual content from four Central and South American countries.



Data Management Components:

- Data description
- Metadata
- Format
- Data organization
- Quality assurance
- Selection and retention periods
- Archiving and Preservation
- Storage and backup/Security
- Existing data
- Access and Sharing
- Audience
- Responsibility
- Intellectual property rights/Copyright
- Ethics and privacy
- Legal requirements
- Budget

In an online search of data management resources, the ICPSR (www.icpsr.umich.edu) performed a gap analysis of known guides from around the world and addressed these similarities and differences in their provided framework for researchers. Given the broad disciplinary scope and comparative work performed by ICPSR, these components were adapted for examining the project and data generated.

Application of Data Management Components

Selected Data Management Components are addressed due to the continuing nature of the project and information availability; the preliminary synopsis of project data presented does not yet include the other steps of the research process, such as analysis, where additional data may be acquired, generated or used. For each Component, general observations and questions are posed to facilitate further discussion on what resources are needed to better facilitate data management.

Data Management Components	
Data description/Metadata/Format	Explanation of data collection process and sources is documented with consistent data types and file formats gathered from each site; what metadata standards(s) could be used to accommodate the diverse data types? Long-term sustainability of acquired formats?
Data organization/Quality assurance	Project data are monitored by a designated research assistant who coordinates with local field sites and is proficient in native languages. Are there 'standard' quality assurance factors that are required for particular data types?
Access and Sharing/Audience	Collected response data are in the regional dialect of country; while translations to English are underway for project partners, how should these different versions be arranged and presented to the greater community?
Archiving and Preservation/Storage and backup	Separate domain-based repository options are available for social and physical science, but for the project as a whole, where would the research data be deposited? Where do you locate this type of information for other countries?

Navigating next steps

Raising awareness of data curation and management within the research process provides an initial step in illuminating what decisions should be considered as data collection progresses. However, additional guidance is still needed to assist researchers:

- Foreign language content, especially data collected from national government agencies, may be privy to additional policies regarding attribution; may require contact with each country's respective agency to resolve.
- Bridging curation resources, such as metadata standards, and calibrating their use and adoption across disciplines would be particularly helpful in approaching the different dimensions of interdisciplinary research.

It is through analysis of diverse research data collections cases that the robustness of data management guides and solutions can be more rigorously assessed.