

INTRODUCTION

A post-doctoral CLIR fellow, Ting Wang, has been investigating the data management (DM) and curation behaviors and needs of faculty members in Lehigh University's Earth and Environmental Sciences (EES) department. Many EES faculty members have participated in detailed interviews as part of this investigation. The interviews have revealed a variety of data-related challenges. Some can be readily addressed by Lehigh's Library and Technology Services, while others pose new challenges. The interviewing process provides a model for similar initiatives at other university earth science programs, as well as a potential model for exploring data needs in other research venues and subject areas.

WORKFLOW

In our assessment of data needs, we are interviewing faculty, researchers, and graduate students rather than conducting surveys. The department size lends itself to intensive interviews that provide a richness of detail that a survey can lack.

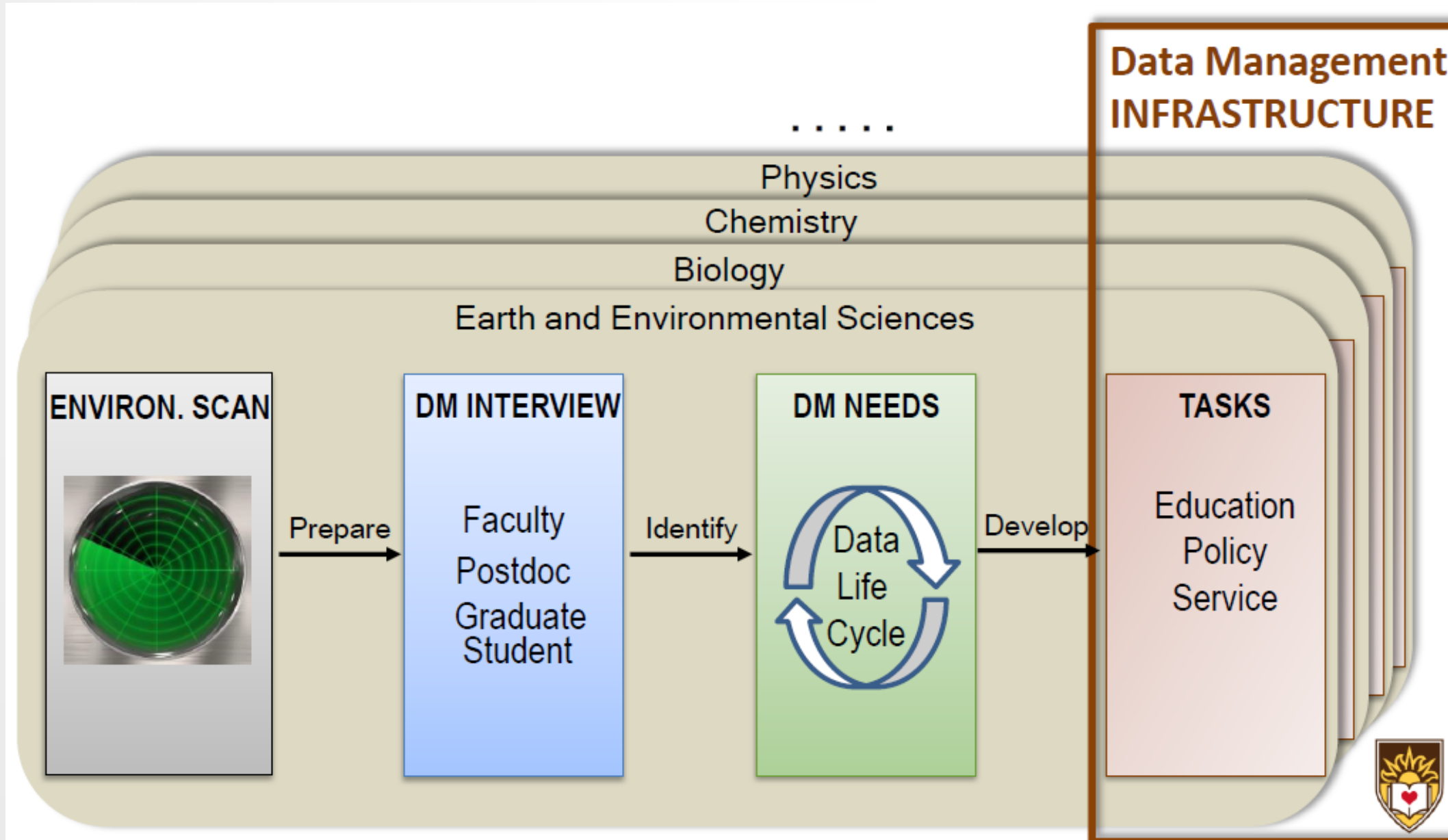


Figure 2. The workflow for our data management interviews in EES and our future plans to achieve the long-term goal, a data management infrastructure for Lehigh University. The "Environmental Scan" on the left refers to our initial phase of reviewing literature about general issues in data management as well as discussing with the EES department chair the research programs and data needs of EES faculty.

QUESTIONNAIRE

We developed an interview questionnaire by referring to the NSF Data Management requirements and guidelines^{4,5,7}, several data management plan (DMP) templates created by the University of Virginia⁶, and the California DMP Tool³. Our questionnaire covers all the relevant DM topics, including types of data, data and metadata standards, policies for access and sharing, policies and provisions for re-use and re-distribution, and plans and provisions for long-term preservation. This questionnaire can also be used when researchers request assistance in DMP preparation.

Before each individual interview, we customize the questions in the questionnaire to the data characteristics of the interviewee's research.

CASE STUDIES

► Interview I

Date life cycle:

The researchers collect rock samples, measure orientation parameters during field sampling, make records in field notebooks, measure the magnetic intensities of samples in the laboratory, analyze the orientation parameters, determine the geographic orientation of the magnetism measured in the laboratory, publish and share research findings at conferences and journals, and archive the full range of data (raw level to intermediate to final products) in lab computers and other storage media.

DM needs:

- Systematically archive all data files generated.
- Preservation of field notebooks.
- Space for physical samples.
- Data quality control to ensure no mistakes are made during data recording or format conversion.

Near-term tasks:

- Create archive system and associated metadata for research group data retrieval.
- Develop metadata for field notebooks and digitize latter for placement in Lehigh's institutional repository.
- Determine disposal policy for useless physical samples.

► Interview II

Date life cycle:

The researchers collect lake sediment cores, measure their lithological and isotopic properties, save the data in spreadsheets, use software to analyze and display the data, present the research and supporting data at conferences and in journals, and archive the full range of data (raw to processed) in lab computers and other storage media.

DM needs:

- Storage and sharing of raw data and complicated datasets.
- Management of physical samples.

Near-term tasks:

- Work underway with IT team to create a website that enables storage and international sharing of data common to the researchers in this specific research area.
- Explore SESAR (System for Earth Sample Registration) as possible system for archiving physical samples.

ACKNOWLEDGEMENTS

Work on this project was done under the auspices of the Council for Library and Information Resources (CLIR) and with active participation of the department of Earth and Environmental Sciences of Lehigh University.

DM BEHAVIORS

The DDI Structural Reform Group¹ and the Oregon State University DM guide² inspired a simple graphical model that illustrates the major stages in the data life cycle for sub-disciplines represented in EES's research:

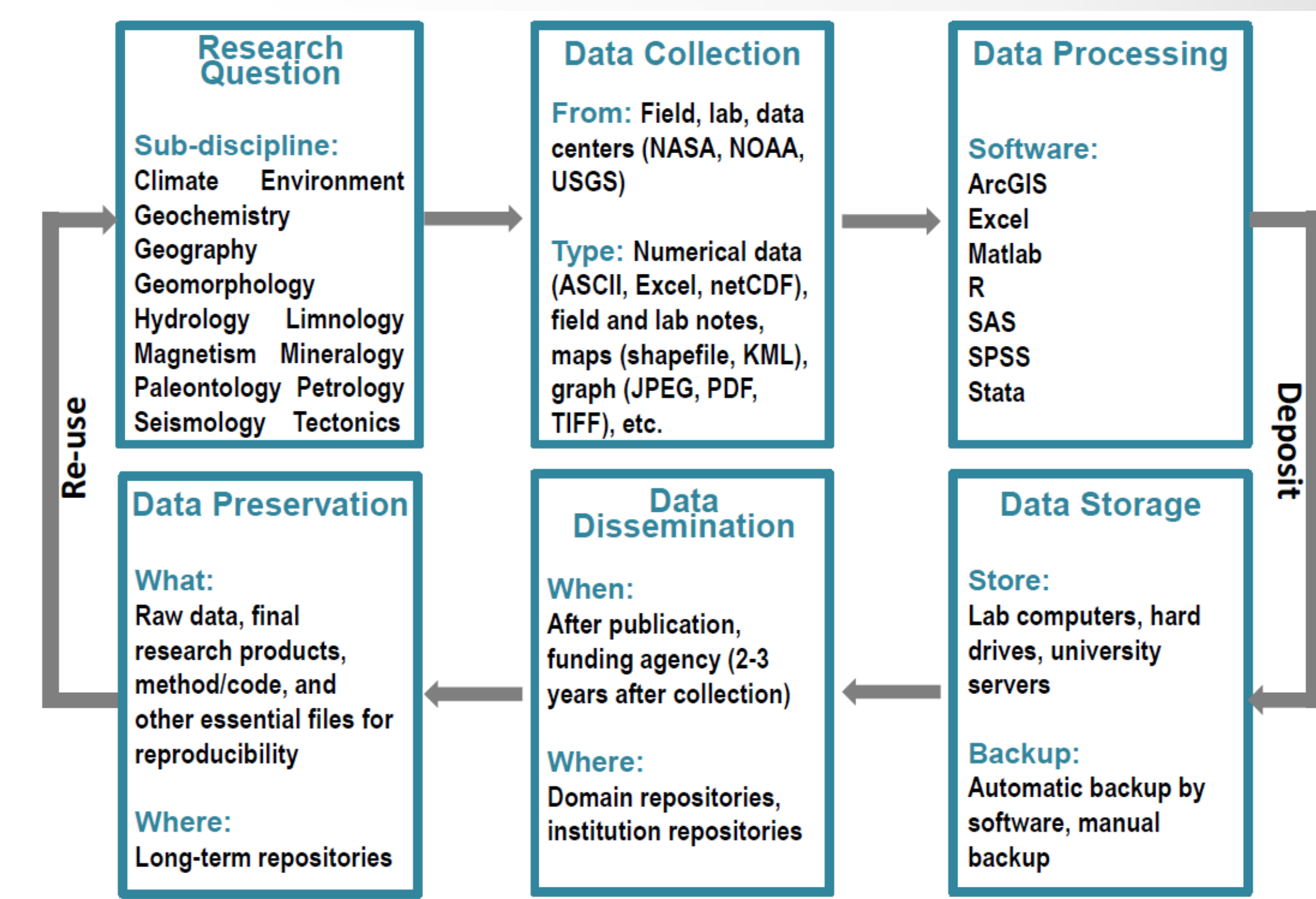


Figure 1. The data life cycle of EES research and summary of data curation tools and behaviors of current EES faculty.

SUMMARY OF DM NEEDS

Our interviews thus far suggest that DM needs in EES across stages in the data life cycle are as follows: (1) Collection: quality control, documentation, metadata standards; (2) Processing: workflow tracking, naming convention, data visualization, and software version compatibility; (3) Storage: space limits, security, and archiving of physical samples; (4) Dissemination: ethical and legal issues, when and where to share; (5) Preservation: data selection, format migration.

FUTURE WORK

We will use the interview framework described above to assess and to help meet data needs in other departments in the sciences as well as in the social sciences.

REFERENCES

1. DDI Version 3.0 Conceptual Model. DDI Alliance, 2004, <http://libraries.mit.edu/guides/subjects/data-management/cycle.html>
2. Data Management & Planning, Oregon State University, <http://ica.library.oregonstate.edu/subject-guide/1346-Data-Management-Planning?tab=515626>
3. DMPTool, University of California Curation Center, <https://dmp.cdlib.org/>
4. NSF Data Sharing Policy at Award & Administration Guide (AAG) Chapter VI.D.4., http://www.nsf.gov/pubs/policydocs/pappguide/nsf13001/aag_6.jsp#VID4
5. NSF Data Management Plan Requirements at Grant Proposal Guide (GPG) Chapter II.C.2.j., http://www.nsf.gov/pubs/policydocs/pappguide/nsf13001/gpg_2.jsp#dmp
6. NSF Data Management Template for Division of Earth Sciences (EAR), University of Virginia Library, <http://www2.lib.virginia.edu/brown/data/NSFDMP.html>
7. NSF GEO Division Data Policies and Information, <http://www.nsf.gov/geo/geo-data-policies/index.jsp>