Cultivating Data Expertise: Comparison of Approaches in Geoscience Data Centers and Academic Libraries

Cheryl A. Thompson, School of Information Sciences, University of Illinois at Urbana-Champaign

Introduction
Data-intensive research promises advancements in knowledge but requires professionals with knowledge in working with research data. As institutions extend their services to data, an understanding of how to build data expertise into service models and staff is needed. This study examines how a mature data center, National Center for Atmospheric Research (NCAR), and exemplar academic library, Purdue University Libraries, develop data expertise and services.

Method
A case study approach is employed to examine data expertise through a series of qualitative interviews and organizational documents. The concept, knowing in practice, developed by Orlikowski (2002) guides the understanding of expertise and competence needed for data work.

NCAR Case
• 18 Interviews with data service managers and staff
• Conducted Nov 2014 – Oct 2015

Purdue Case
• 12 Interviews with data librarians and campus partners
• Conducted Aug – Sep 2015

Analysis
• Qualitative codes develop inductively and deductively.
• Analysis using ATLAS.ti for Mac.

Participant demographics
NCAR (N=18)
• Average time in position: 12.4 years (min=0.5, max=30)
• Positions titles includes Project Scientist, Database Engineer, & Software Engineer

Purdue (N=12)
• Average time in position: 7.91 years (min=0.5, max=15)
• Positions titles includes Data Specialist, Data Curator, and Liaison Librarian.

Acknowledgments: This research was supported by the Data Curation Education in Research Centers (DCERC) program, funded by the Institute of Museum and Library Services (RE-02-10-0004-10).

Data Expertise

Core Knowledge
• Preserving data
• Organizing scientific data and collections
• Standardizing metadata, data, and approaches
• Understanding use and users
• Understanding research process
• Harnessing technology for data solutions
• Leading services

Data service roles
Data service managers
• oversight for data services/units.
• responsible for service activities and performance.

Data curators
• responsible for the curation of research data.
• primarily concerned with ensuring reuse.

Data engineers
• building software, automating processes, and computational solutions for data.

Data consultants
• provide guidance, best practices, and resources for usersthrough consultations and trainings.

Data Center – NCAR Case
• Data scientists
• designing simulations and data products
• responsible for data management.

Data Center – Purdue Case
• Academic Library – Purdue Case
• Training users
• Discovering data for users
• Building relationships with community
• Understanding data landscape
• Collaborating with stakeholders
• Understanding data literacy

Key factors to building data expertise
Primary factors:
• Data champions
• Grass roots efforts
• Exploratory projects
• Admin. buy-in
• Retraining staff

Secondary Factors
• Restructuring organization
• Resource allocation
• Partnerships
• Collegiality
• Stakeholder interests

NCAR and Purdue have developed in-house data expertise by embedding knowledge, skills, and roles into their organizational structure. Data professionals offered valuable contributions in the areas of organizing data, understanding data uses and users, harnessing technology for data solutions, and standardization. NCAR staff require expertise in data analysis and common instruments/models, where Purdue embeds skills in building relationships and instruction in their data staff.

New data roles have evolved to embed data expertise into service models. Explicit career paths for data professionals are limited but starting to emerge.

Top level support, champions for data services, and other contributing factors have enabled data expertise and service development in these sites.

Analysis is ongoing. Next steps includes further comparative case analysis how organizational context impacts data expertise and services models.

Conclusions

References

Image Credits
Copyright University Corporation for Atmospheric Research (UCAR). Licensed under a Creative Commons Attribution-Non Commercial 4.0 International (CC BY-NC 4.0) License, via OpenSky.