Data Librarians

Wade Bishop, Associate Professor
@WadeBishopUTK
https://orcid.org/0000-0002-5022-2707
School of Information Sciences
University of Tennessee

Participation Activity #3 discussion

- Introduce the community you chose and why you chose them.
- Were you a member?
- Does this might be a domain, discipline, organization, profession, or other group need librarians?
- What are some roles librarians may serve?

Image Credit: Gerd Altmann from Pixabay

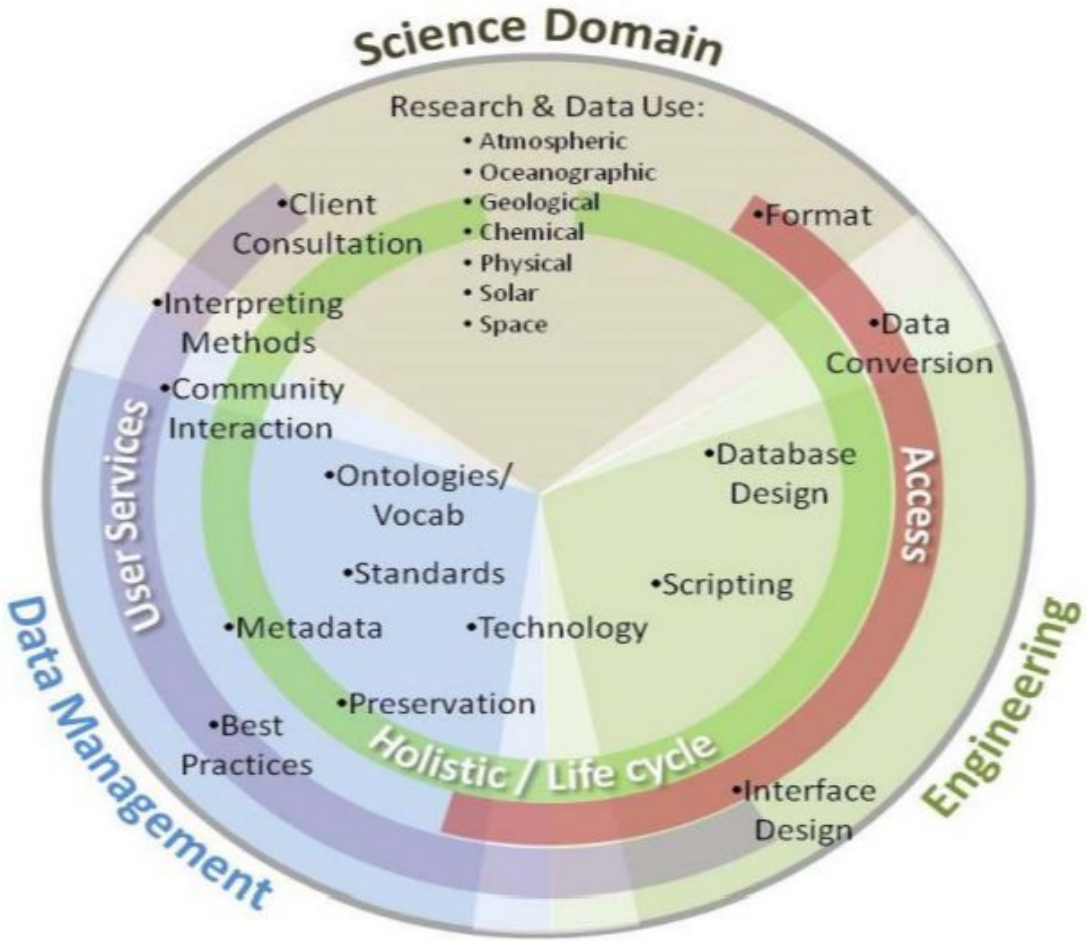
Data librarians

- As research becomes more data-intensive and researchers face new challenges in managing and sharing research data, libraries have begun to offer a variety of data support services, including instruction and training, data management planning guidance, data stewardship and curation, and data visualization.
- Some libraries have tasked liaison librarians with providing data services to their groups as part of their broader charges, while others have hired librarians and information specialists dedicated full-time to providing data services

Federer, 2018

- More than two-thirds of respondents (n=55) supported more than 1 academic discipline, with a mean of 2.8 disciplines supported.
- Only 5% (n=4) reported spending all of their time on data-related work. The majority of respondents (n=48, 59%) reported spending at least 50% of their time on data-related work, with a mean of 55% of time spent on data related work.
- ALA-accredited master's degree: 83% (n=68); undergraduate science degree: 39% (n=32); science master's degree: 23% (n=19); doctorate (PhD) (any discipline): 17% (n=14).
- The top 5 items overall were: "Developing relationships with researchers, faculty, etc."; "Oral communication and presentation skills"; "Teamwork and interpersonal skills"; "Written communication skills"; and "One-on-one consultation or instruction."

Fig. 1. Data curation education skill spectrum



Mayernik, M.S., Davis, L., Kelly, K., Dattore, B., Strand, G., Worley, S., & Marlino, M. (2014). Research center insights into data curation education and curriculum. In Ł. Bolikowski, V. Casarosa, P. Goodale, N. Houssos, P. Manghi, & J. Schirrwagen (Eds.), Theory and Practice of Digital Libraries -- TPDL 2013 Selected Workshops (Vol. 416, pp. 239–248). Springer International Publishing.

Some Challenges and a Need

- These roles for liaison librarians come before and after scientific processes (e.g., data collection and analyses).
- These liaison tasks are seen as non-analytic parts of data science.
- Data producers may have to do liaison and curation tasks without the range of skills, knowledge, and abilities to do so effectively.
- Some doing this work may not even consider these tasks work (or consider there are other/better ways to do them).
- Still, the shortage of trained information professionals compared to the amount of data curation activities needed is well known and several efforts have occurred to address this shortage.

Developing a Curriculum

- The Developing a Curriculum (DACUM) approach creates a list of knowledge, skills, and abilities, operationalized job descriptions, and eventually learning outcomes for use in continuing education.
- The DACUM approach assumes job incumbents know their job *better than anyone else*, and therefore they are the best at describing what it is that they do.
- The best way to define a job is by describing the specific tasks that are performed on the job. Like any other intellectual work, those employed in information agencies routinely have tasks that may be difficult to describe.

DACUM approach

- The DACUM method offers advantages over other job analyses because of a relatively lower cost and time commitment (e.g., direct observation).
- A researcher should eliminate bias and sample evenly across a profession.

1. Interviews with job incumbents
2. Survey Validation with broader workforce

Methods

- Recruitment was conducted by snowball sampling at a national conference on earth science data with the incentive of helping create a core of expertise for earth science data management.
- All twelve participants managed earth science data as at least part of their position, but for most it was their entire job.
- The interview schedule consisted of ten questions, gathering information on:
  1) current job title and work setting;
  2) total years working in current job and with earth science data;
  3) credentials, degrees and other applicable education or training;
  4) daily, weekly and less frequent tasks associated with job; and
  5) tasks not associated with the job.
- Interviews were recorded, transcribed, and coded in NVivo.
- Emphasis was placed in identifying themes relating to the DCC model's sequential lifecycle actions.

Bishop, B. W. & Hank, C. F. (2018). Earth Science Data Management: Mapping Actual Tasks to Conceptual Actions in the Curation Lifecycle Model. In: Chowdhury G., McLeod J., Gillet V., Willett P. (Eds.). Transforming Digital Worlds. iConference 2018. Lecture Notes in Computer Science, Vol. 10766. Springer, Cham. https://doi.org/10.1007/978-3-319-78105-1_67

Job Analysis – Earth science librarians & data managers

- First CALL job analysis
  - Learning objective 1
- Led by Ashley, Jan-Apr 2020
- Recruitment through ASLI, ESIP, WHOI
  - In-person and email
- 12 interviews
  - 6 librarians and 6 data managers

Bishop, B. W., Orehek, A. M., & Collier, H. R. (under review). Job analyses of earth science data librarians and data managers. *Bulletin of the American Meteorological Society*.
Orehek, A. M., Bishop, B. W., & Collier, H. R. (under review).
What (atmospheric) science librarians really do? *2021 AMS Annual Meeting*.

Background and Experience

- Even distribution
  - 6 w/ science undergrad, 6 w/o; 6 w/ MLIS-related, 6 w/o; 3 PhDs
- Variety of professional certifications
- Variety of previous experience
  - Libraries, labs, repositories/DBM
- Average years in field – 16
- Year range in current job – 2-16
- Skills and tasks varied
  - Mostly OTJ or self-taught

Skills and Tasks

Skill sets

- Programming languages
  - Python, Latex, MATLAB
  - R, SQL, SQLite, SPARQL
- Software and frameworks
  - ArcGIS, QGIS GeoServer
  - Cyberduck
- Data carpentry
- Metadata
  - Open Refine, XML editor, Intergraph metadata
- Information modeling

Job Tasks

- Communication
  - Most common theme, Emphasis on soft skills overall
  - “Translator” between researcher and user
- Collaboration
  - Common theme, Daily/weekly/less frequent
- Outreach
- Teaching & research
  - Tenure-track, Enrichment projects
  - Workshops, Classroom visits
- Data-specific tasks
  - Data storage, DMPs, Repositories

Takeaways

- Job titles meaningless
  - Tasks and titles don’t always relate, Ambiguous
  - Constantly evolving fields
- Varied education
  - STEM field familiarity and social norms
  - LIS hones DBM and other information skills
- Communication important
  - Huge collaborative effort by all parties
  - Efficiency and effectiveness throughout data life cycle

Job Posting Analysis Paper

Methods:

- Selected job postings from Indeed.com for different data manager and data librarian positions.
  - Tried to get representative positions from various types of organizations.
- Coded each posting in NVivo for different skills required, common tasks for the position, and experience.
- Created tables and developed results.

"…librarians tend to need previous library experience and general knowledge of library systems, while data managers often need more technical abilities within data management. Research experience is important for both."

Student perspective | Interviewing

- Persistence!
- Ask questions in order and verbatim for consistency
  - Easier to code later
- Clarify unknown words or acronyms used by interviewee
- Take notes during interview, helps clarify phrases when transcribing and capture interviewer thoughts
- If you never interviewed/transcribed before, listen/read through an example from PI
- Anonymize all names and locations in transcriptions

Student perspective | Coding

- Self-taught on NVivo via YouTube
  - LinkedIn Learning and OIT workshops also good resources
- Read through transcriptions before starting, if you didn't transcribe them
- Collaborate on a codebook, if working with a partner

CALL Year 2 Activities

- Original research on science liaison librarianship job analyses
  - Interviewing job incumbents—:
    - Bishop, Bishop, B. W., Smith, P., Eaker, C., & Orehek, A. M. Data librarian duties: Job analyses and perspectives on data management plan compliance and evaluation. *Journal of eScience Librarianship*
    - Other liaison librarians? data managers?
- Survey validation?
    - CALL survey of data librarians, earth science data managers, and others?
- Job descriptions?

Survey validation

- The survey validation approach has been used by a variety of other professions such as teaching, nursing, and law to inform curricula with empirical data from actual practice.
- The curriculum development plan for the Geographic Information Librarianship (GIL) project includes a survey of practicing GIS and map librarians, archivists, and other information professionals to weigh the importance of listed knowledge, skills, and abilities (KSAs) from their real-world experience to inform curriculum development.
- Now that we know typical tasks... we can do this for ESIP and likely others?

Resources

- Federer, L. (2018). Defining data librarianship: a survey of competencies, skills, and training. *Journal of the Medical Library Association: JMLA*, 106(3), 294. 10.5195/jmla.2018.306