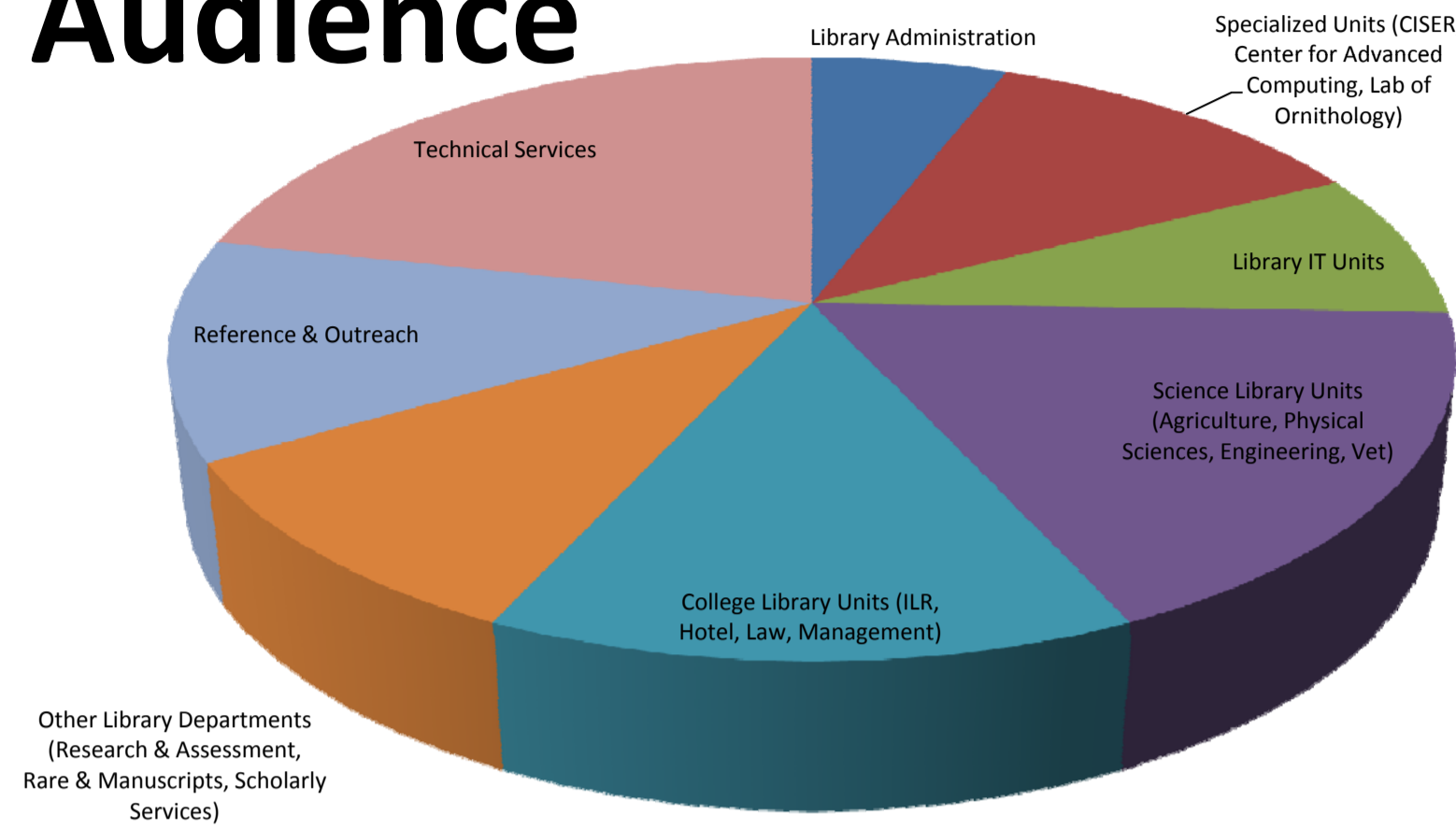


The Data Discussion Group: Growing a data-curation aware community of librarians

Dianne Dietrich, Keith Jenkins, Gail Steinhart - Cornell University Library

Audience



Feedback

Most Helpful/Relevant Sessions

Reference services issues
Finding/using free data sources
Census, UN, Government Data
Open access issues

Less Helpful/Relevant

Data curation issues
Talking to faculty
Cyberinfrastructure
Roles for libraries

Building awareness of data curation issues can come via discussion of issues using secondary data, including metadata and discoverability, practicality of certain file formats, and restrictions and licensing. The positive response to the meeting featuring Open Access issues suggests this is a viable strategy.

What's helped

"Learning about the new mandates and initiatives has been extremely helpful."
"Awareness of new data sources."
"I like the mix of project and resources topics."

Suggestions for improvement

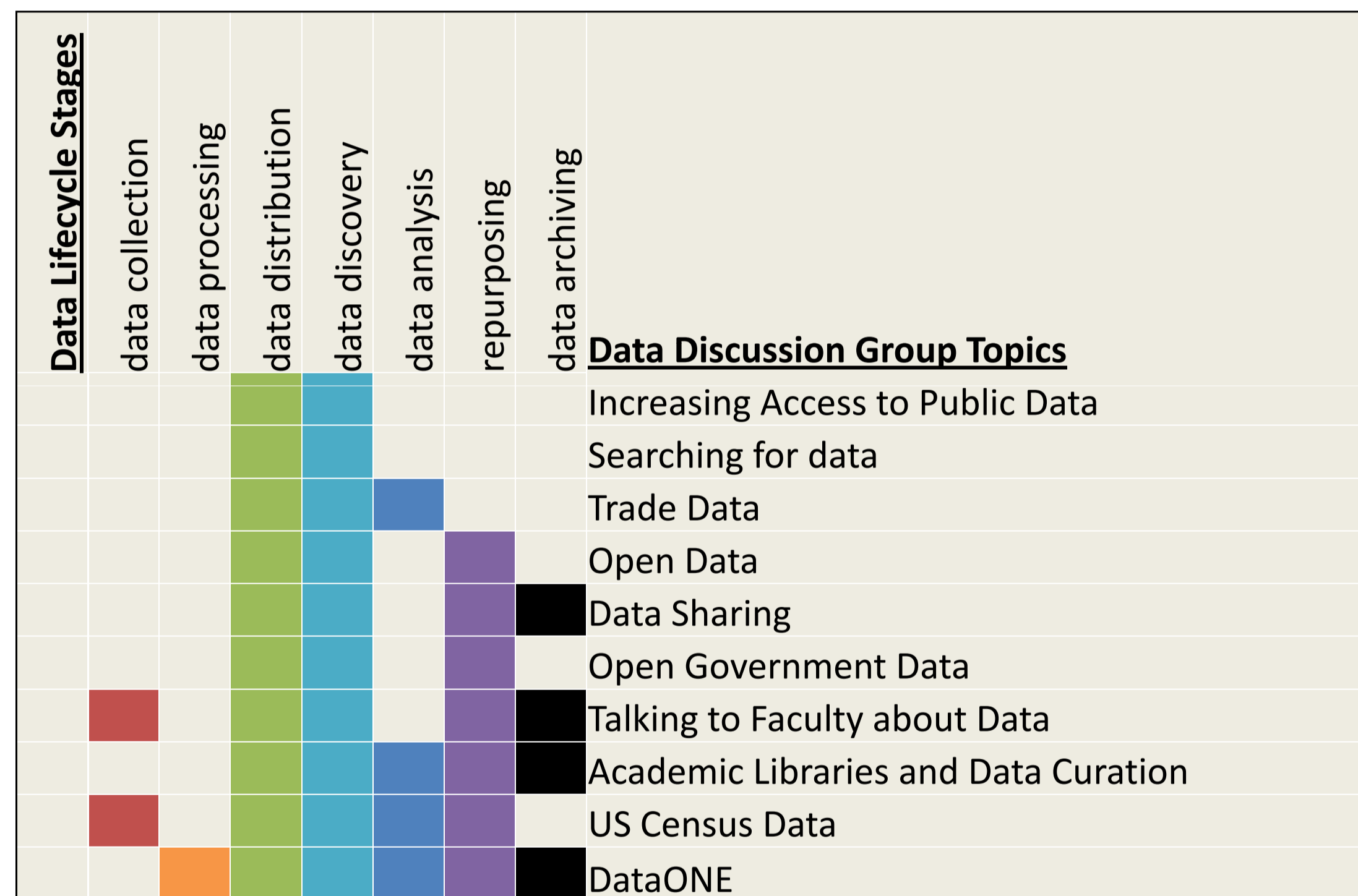
"I'd like to see more graduate students and/or faculty attend it in order to get different perspectives."
"Perhaps some researchers talking about their data - how they collect, transform, publish."
"Maybe this group could lead [an] effort [to devise a good system for discovering subscription and public data sources]."

Expanding to include a wider range of perspectives – such as how faculty and researchers use data – would be helpful. The group's greatest opportunity for growth appears to be exposing library staff to a wide range of data issues.

Abstract

The Data Discussion Group at Cornell University Library gathers monthly to discuss a wide range of issues relating to data curation and the library. The group emerged as the result of a recommendation in the Data Working Group whitepaper to offer professional development activities for staff at the library in the areas of cyberscholarship, data-driven research, and data curation. The group has completed its first year of programming, and we believe its model may be of interest to others who are interested in creating professional development opportunities for libraries in the area of data curation.

2009-2010 Programming Themes



* Lifecycle based on DDI Conceptual Model, <http://www.iasistdata.org/blog/conceptualizing-digital-life-cycle>

Gaps

Few topics will touch on all stages of the data lifecycle, so it is important that the group consider how best to distribute the meetings to ensure that a range of data issues from throughout the lifecycle is presented.

Several meetings have touched on data collection and data processing, suggesting an opportunity to address these data issues. When we consider this gap with our survey feedback (left panel), this indicates that it might be helpful to invite faculty to future meetings to talk more about how they create and use data in their research.

Contact information:

{dd388,kgj2,gss1}@cornell.edu

Takeaways

Keep it fairly informal

The majority of our meetings were not very structured; the coordinators prepped themselves with only a few talking points and ideas to facilitate the conversation. This encouraged a participatory environment among the attendees, especially those who didn't consider themselves well-versed in "data curation" issues.

Connect data curation to established library activities

Librarians have been helping patrons locate, and in some areas, use secondary data for quite some time. Many of the issues reference librarians face when helping patrons connect to issues of data curation – for instance, understanding why a certain data source is hard to use due to legacy file formats, can prompt a discussion on the importance of choosing preservation-friendly formats when archiving data.

Encourage a range of perspectives

Consider and invite other groups within and beyond the library that work on issues related to data curation. For instance, archivists are well-acquainted with issues of selection, appraisal, and long-term storage of unique records: this perspective is extremely valuable for librarians doing data curation. Technical services and metadata professionals understand how description can affect discoverability of resources; other individuals can bring to the table an understanding of technical challenges to working with digital data.

Vary the meeting format

We found that attendees responded positively to a flexible meeting format. Some meetings were more like discussion groups; other times, we had presentations from various individuals; one meeting was structured as a journal club. Since data curation is a new realm for many librarians, we believe that offering a variety of formats to participate in the group helped us to be more inclusive and accessible.