Table of Contents

THE QUESTIONNAIRE

CHAPTER 1 – Introduction

1. Please provide the following contact information:

    A. Name:
    B. Organization:
    C. Work Title:
    D. Country:
    E. Email Address:

2. Your college is:

    A. Public
    B. Private

3. The Carnegie Class of your college might best be described as:

    A. Community or 4year
    B. College
    C. MA/PHD level college
    D. Research University

4. The full time equivalent enrollment for your college for all schools and departments is approximately:

5. Which 3 subject areas account for the largest share of the large scale data sets that your library handles?

6. As a rough estimate what percentage of the total data in your data repository are accounted by each of the top three subject areas? Please list the subject areas and your estimate:

    A. Subject area 1
    B. Subject area 2
    C. Subject area 3

7. Does the library offer to faculty advice on how to develop data management plans for grant proposals or personal use?

    A. Yes
    B. No

8. Does the library have any series of tutorials, seminars, classes, videos or other formal means to train faculty in data management?

    A. One on one tutorials
    B. Formal classes
    C. Workshops or seminars
    D. Videos or web based tutorials

CHAPTER 2 – Resources for Data Curation

9. What percentage of the college/university overall spending on data curation would you say is contributed by the library?

10. Has the library ever received an external grant to support activities in data curation?

    A. Yes
    B. No

11. If so what is the sum total of the amount your library has ever received to develop or support data curation efforts?

CHAPTER 3 – Cooperation Among Departments

12. Describe your library's relations with other players in data curation such as the College/University office of grants management, or the academic departments that use the data curation services? Do they understand the library's efforts? Are there turf battles? What is the overall level of cooperation?

13. Since many grants now require a data curation strategy, and many departments of a college may be involved in this effort, the cost of this effort should be shared among these departmental players. Is this the case at your institution? If so, how do the contributions and roles break down? Has the library been given extra funds to carry out its data curation role?

14. Which academic fields would you say take most advantage of the data curation services offered by the library:

15. Which phrase do you feel best summarizes your institution's approach to data management and data curation? (Choose the phrase which is more true than the alternative.)

A. We are developing a stand alone data repository that serves the unique needs of large scale data projects and management decisions are in the hands of experienced data curators.
B. We are closely integrating our data repository with our pre-existing institutional digital repository and though we have experienced data curators management control ultimately rests with the library committees on open access and repositories.

CHAPTER 4 – Data Security

16. Describe the system that your library uses in determining the proper level and timing of access for data sets maintained by the library or other archival custodian. For example, which are available for public access, which only for scientists working on a particular project and which only after an embargo period after publication of a journal article based on the data. The possibilities for significant mistakes are ample; how does your organization avoid them?

CHAPTER 5 – Metadata and Citation

17. What percentage of the data sets contributed to your repository come with pre-existing metadata developed by the originator of the data?

18. Describe some of the practices that your library follows in developing metadata practices for the proper classification and recall of data sets?

19. How easy has it been to develop metadata for following types of data:

    A. Notes or logs from scientific/social science experiments
        i. Quite easy
        ii. Relatively easy
        iii. Able to do with modest challenges
        iv. Relatively difficult
        v. Very difficult

    B. Output or results from medical or scientific instruments or other monitors
        i. Quite easy
        ii. Relatively easy
        iii. Able to do with modest challenges
        iv. Relatively difficult
        v. Very difficult

C. Video, photographs or other images
   i. Quite easy
   ii. Relatively easy
   iii. Able to do with modest challenges
   iv. Relatively difficult
   v. Very difficult

D. How easy has it been to develop metadata for spreadsheets?
   i. Quite easy
   ii. Relatively easy
   iii. Able to do with modest challenges
   iv. Relatively difficult
   v. Very difficult

E. Databases
   i. Quite easy
   ii. Relatively easy
   iii. Able to do with modest challenges
   iv. Relatively difficult
   v. Very difficult

F. Software code
   i. Quite easy
   ii. Relatively easy
   iii. Able to do with modest challenges
   iv. Relatively difficult
   v. Very difficult

20. About what percentage of the total staff time that your library spends on data management issues would you say is accounted for by the development of metadata and other cataloging issues?

CHAPTER 6 – Personnel

21. Has the library hired or assigned librarians specifically to work on data curation issues?

    A. Yes
    B. No

22. Does your library have a "data curation" staff assigned specifically to deal with this issue?

A. Yes
B. No

23. If yes how many FTE positions have been assigned to this staff or department?

24. Is there a distinct line item in your library budget for data curation?

    A. Yes
    B. No

25. Do you believe that recent MLS graduates hired by your library have been adequately versed in data curation issues?

    A. Yes
    B. No

26. What is the total number of grant proposals that have been reviewed or contributed to by your data curation office or service in the past year.

CHAPTER 7 – Interdepartmental Cooperation

27. Data curation issues are often dealt with by interdepartmental teams from the library, the office of research grants management, information technology and researchers themselves. Has your college established such an interdepartmental team?

    A. Yes
    B. No
    C. Not formally but perhaps informally

28. What are the respective roles at your institution of the library, the office or research management, information technology, and the researchers themselves in data management and preservation issues? Has there been general agreement on the roles of each party? Have other departments played a prominent role and if so, what?

29. Studies have shown that centralized data repositories in data intensive fields can significantly improve scholarly productivity by dramatically reducing the amount of time it takes to locate important data sets. However, the costs of discovering, centralizing, cataloging, storing, distributing and even marketing centralized data repositories can be considerable in terms of software, computing costs and particularly personnel and training costs. How are these costs determined and

borne by various institutional players at your institution?

CHAPTER 8 – Data Integrity

30. Have you had at your institution any data curation disasters in which critical data was lost, destroyed, impermissibly accessed or otherwise rendered far less useful than it might have been with more adequate curation measures. Keeping in mind that responses are not connected to institutions or respondents can you describe some of these instances?

31. Does the library have a data curation data preservation strategy through which it has decided to maintain data for a minimum number of years?

    A. Yes
    B. No

32. If yes what is the minimum number of years?

33. How easy has it been to procure and archive the following types of data?

    A. Notes or logs from scientific/social science experiments
        i. Quite easy
        ii. Relatively easy
        iii. Able to do with modest challenges
        iv. Relatively difficult
        v. Very difficult

    B. Output or results from medical or scientific instruments or other monitors
        i. Quite easy
        ii. Relatively easy
        iii. Able to do with modest challenges
        iv. Relatively difficult
        v. Very difficult

    C. Video, photographs or other images
        i. Quite easy
        ii. Relatively easy
        iii. Able to do with modest challenges
        iv. Relatively difficult
        v. Very difficult

D. How easy has it been to develop metadata for spreadsheets?
   i. Quite easy
   ii. Relatively easy
   iii. Able to do with modest challenges
   iv. Relatively difficult
   v. Very difficult

E. Databases
   i. Quite easy
   ii. Relatively easy
   iii. Able to do with modest challenges
   iv. Relatively difficult
   v. Very difficult

F. Software code
   i. Quite easy
   ii. Relatively easy
   iii. Able to do with modest challenges
   iv. Relatively difficult
   v. Very difficult

CHAPTER 9 – Assessment

34. How have your library's data curation activities impacted the standing and scholarly results of your institution? Have you made efforts to assess the impact of the data curation activities on institutional success?

CHAPTER 10 – Information Resources for Data Curation

35. What are some of the data archives or directories that you have found most useful in your data curation pursuits?

36. Which blogs, listservs, websites, magazines, newsletters, conferences and other information resources have you found most useful in your data curation activities?

37. Which institutions do you most admire for their data curation efforts?

SURVEY PARTICIPANTS

Binghamton University Libraries
Central Washington University
Colorado State University
Columbia University
MIT
Montana State University
Northern Illinois University
Roskilde University
Stellenbosch University
University of Brasilia
University of Nebraska-Lincoln
University of Nevada Las Vegas
University of Surrey
University of Toronto
Virginia Commonwealth University