Demystifying the Data Interview

Eugenia S. Kim
_Purdue University_, eugenia.kim24@gmail.com

Jake R. Carlson
_Purdue University_, jrcarlso@purdue.edu

Follow this and additional works at: http://docs.lib.purdue.edu/lib_research
Part of the Library and Information Science Commons

http://docs.lib.purdue.edu/lib_research/126

This document has been made available through Purdue e-Pubs, a service of the Purdue University Libraries. Please contact epubs@purdue.edu for additional information.
**QUESTION**
What would a typical librarian need to know ahead of time in conducting an interview with a faculty researcher regarding his/her research data and associated needs in managing or curating this data?

**METHODS**
- The authors analyzed the DCP Toolkit to identify specific definitions, concepts, resources, and examples to provide a basic understanding of data curation.
- By compiling this information with additional resources for reference, the authors intend to help librarians take full advantage of the DCP Toolkit.

**ABOUT THE DATA CURATION PROFILES**
- From 2007 - 2010, the Purdue University Libraries and the Library School at the University of Illinois conducted research on “Which researchers are willing to share data, when, with whom and under what conditions?”
- Research was supported by the Institute of Museum and Library Services.
- The results led to the creation of the Data Curation Profiles (DCP) Toolkit.
- The DCP Toolkit is designed to assist librarians in identifying the data management needs of researchers.

**RESULTS**
- Data concepts and terminology and meaning are often vary between various fields.
- Concepts do not necessarily have examples that can be easily referenced by librarians.
- These results informed the development of the Data Curation Toolkit Workshops.

**CONCLUSION**
- The results of this project will be further analyzed through:
  - Assessing the impact of the training delivered in the DCP workshops.
  - The effect of the workshop and DCP toolkit in enabling librarians to conduct data interviews.
  - Other sources such as the community forums on the DCP Toolkit website.

**RECOMMENDED RESOURCES**
- Data Curation Profiles Community [http://www.lib.purdue.edu/dcp/](http://www.lib.purdue.edu/dcp/)
- Resources for Digital Curators | Digital Curation Centre [http://www.dcc.ac.uk/resources/](http://www.dcc.ac.uk/resources/)
- Digital Curation | ICPSR (Inter-University Consortium for Political and Social Research) [http://www.icpsr.umich.edu/icpsrweb/ICPSR/](http://www.icpsr.umich.edu/icpsrweb/ICPSR/)

---

### Demystifying the Data Interview

**Eugenia Kim** - Data Curation Intern, Purdue University Libraries - eugenia.kim24@gmail.com

**Jake Carlson** - Data Research Scientist - Purdue University Libraries / Distributed Data Curation Center - jrcarlso@purdue.edu

---

### Data

**Purpose:** A clear definition for “data” is essential for setting the context of a data interview.

**Challenges:** The term “data” is often defined very broadly. It may mean different things to different people. Some disciplines, such as the Humanities, may not think in terms of “data”.

**Examples of Data:** notebooks, data files, images, e-mails, surveys, videos, transcripts, drafts of publications.

**Selected Definition:** “Research data is defined as the recorded factual material commonly accepted in the scientific community as necessary to validate research findings” - OMB Circular A-110.

**Source:** [http://www.whitehouse.gov/omb/circulars_a110](http://www.whitehouse.gov/omb/circulars_a110).

### Data Curation

**Purpose:** To understand or be familiar with the tasks and responsibilities that are generally associated with the term ‘data curation’. Specific activities associated with curation are “archiving” and “preservation”.

**Challenges:** Not every dataset follows the same lifecycle or understands curation in the same way. A librarian may also find that a researcher asks for a task to be performed after the interview or has an unusual type of need.

**Selected Definition:** “The activity of managing and promoting the use of data from its point of creation, to ensure it is fit for contemporary purpose and available for discovery and re-use.” - Lord, Macdonald, Lyon & Garetta (2004) “From data deluge to data curation.” Proceedings of the UK e-Science All Hands Meeting 2004, 31st August - 3rd September, Nottingham UK.

### Data Librarian

**Purpose:** Librarians need to be able to explore how their knowledge, skills and perspective could be applied towards supporting researchers’ needs without feeling overwhelmed.

**Challenges:** Data librarians may come from all kinds of backgrounds and may be called upon to fulfill a wide range of roles requiring a variety of skillsets. Furthermore, the position of “data librarian” has not yet been established at many institutions or may be tied to traditional library services rather than curation.

**Base Identity:** Broadly defined, a librarian with responsibilities relating to the curation and maintenance of data.

---