IFLA/ALA Webinar – Big Data: New Roles and Opportunities for New Librarians

Data Visualizations Meet Libraries: New Opportunities

Dr. Michelle Chen
School of Information
San José State University

June 15, 2016
The “Big Data” Era: How to prepare, how to survive

• What are the implications of big data?

• What is the best way to organize data so that it can be easily found?

• What is the role of the information professional in this digital universe?
The Four V’s of Big Data

• Volume
• Velocity
• Variety
• Veracity

Image Credit: New York Public Library
What’s Happening in LIS?

- The policy landscape
  - NIH-funded results to be publicly deposited (PubMed Central)

  **When and How to Comply**

  1. Preparing a manuscript
  2. Accepted for publication
  3. Reporting to NIH

  **Overview:**
  To advance science and improve human health, NIH makes the peer-reviewed articles it funds publicly available on PubMed Central. The NIH public access policy requires scientists to submit final peer-reviewed journal manuscripts that arose from NIH funds to PubMed Central immediately upon acceptance for publication. [more]

  Show me specific instructions for my publication

  Video Training

- Data is getting bigger and bigger
  - Example: Harvard released big data for books
    - 12 million books, videos, images, maps, etc.
    - From 73 libraries
    - > 100 different attributes for a single object

Image Credit: Digital Public Library of America
What Can We Do?

• What is an effective way to organize, manage and present data?

Data Visualization

Kindred Britain, Stanford University Libraries
What Has Been Done? (Some Examples)

• Health Science Library, University of North Carolina at Chapel Hill
• Renaissance Computing Institute

• Major visualization infrastructures were deployed
• Visualization applications were developed

What Has Been Done? (Some Examples) (cont.)

• Seattle Central Library

  • “Making Visible the Invisible” project, 2005-2014
  • 6 LCD screens on glass wall, 45” x 24”

Image Credit: http://www.mat.ucsb.edu/g.legrady/gWeb/Projects/spl/spl.html
What Has Been Done? (Some Examples) (cont.)

• Citation visualization

Microsoft Visual Explorer

Web of Science
New Opportunities

- Data visualization for better data discovery

- Library is a data center! (e.g., usage/circulation statistics, collections and archives, etc.) E.g., California Library Statistics

- Discover trends, correlations, patterns, and exceptions

→ Better services and operations!
New Opportunities (cont.)

• Data visualization for better data communications
  • Promote information literacy or library events
  • Enhance user experiences (e.g., “AquaBrowser” by ProQuest)

• Governance and fund seeking

→ Better exhibits and persuasiveness!
New Opportunities (cont.)

- Understanding data for fulfilling social responsibility role
  
- Enhance the library’s role in addressing poverty and illiteracy (Maret 2011*)

→ **Improve the well-being of our communities!**

---

Getting Prepared

• Be familiar with data intensive-related issues and problems

• Gain knowledge and skills in handling large-scale data sets

• Be able to understand, interpret, and communicate large-scale data sets

• Participate in seminars, webinars, MOOCs, etc. to gain up-to-date information