Breakout Session Ideas

NB
Feel free to add topics, broad areas of exploration (h3 heading), or expression of interest (use a “+”). Use the space below “Scratchpad” to capture any remaining thoughts or ideas that don’t seem to fit in the session proposals. Thanks!!

Python-Related
• Python for complete beginners ++
• Useful python libraries for working with data +++
• NLTK (Natural Language Toolkit) ++
• How do you learn code? / Code Club ++
• …

Machine Learning
• Feature Engineering ++
• Python/R Libraries for ML tasks +
• …

XML/TEI
• Transformations (XSL/JavaScript/Xquery) +++++ --Nathan

Digital Humanities-Related
• Whatever +++
• Acquiring/Displaying/Licensing Textual Data ++
• Text Analysis / Text Mining +++
• R +

Data Wrangling
• Processing data / cleanup ++++ -- Devin
• Traditional methods (coding, statistics, etc) ++
• Working with numeric data
• Google/Open Refine (if anyone know how to use this) +
• Deduping
• Data Accessibility
• Data Reuse
• …

Research Data Related
• Sharing, Publishing and Citing data +
• Intellectual Property & Data ++++
  o Licensing
• Research data services -- brainstorming from the basics to the grandiose +
• Institutional Repositories
• ...

Data Storage
• Backups, Redundancy +
• Tape, Hard Disk and Solid State applications ++++
• Costs, tiers, hardware refresh cycles, local vs cloud+
• Good Enough & Best Practice ++++
• Recovering from failure +
• ...

Types of data
• Circulation, acquisitions, digitization data +++
• HR data +
• Collections/catalog data (“metadata” or otherwise) ++++ --Aaron
• Traditional metadata -- administrative, technical, preservation, descriptive, etc +
• Everything +++
• Spectrum of data value
• Structured +
• Metadata / catalog / xml / EAD / MARC ++++
• Financial data
• Data about users / member data / patrons ++++ -- Justin
  o Community organizations
  o Data has a community
• Records
• Legacy data
• Data vs Information
• Unit / Metric -- tells story
• Research data / evidence
• Text / books
• Audio

Data Visualization
• Guidance on selecting the right visualization ++
• Beginner software +
• Coding & visualization +++
• ...

Other Topics!
• Repository (Hydra) services for presenting/using research data.+
• Using Data / Interfaces and their users+
• EAD

Scratchpad

• general thoughts!
• I’d like to come away from the workshop knowing how to get started cleaning up a data set
• I’m completely new to this so I’d like to leave with some ideas, software, methods, etc. for beginners.
• I’m interested in how libraries/archives can anticipate/facilitate the research data needs of patrons. What can we do better when processing/preparing data? What tools or resources do we need to make available on our website or in the reading room?