Information Visualization Skills for Academic Librarians: A Content Analysis of Publications and Online Resources in the Digital Humanities

Dr. Michelle Chen
May 5th, 2017
Agenda

• Overview

• Literature review

• Methodology

• Results

• Discussions

• Demo
Overview

• Digital humanities: data has become more voluminous, versatile, accessible and digitized

• Examples:
  - Ferguson, 2017
  - Murrieta-Flores et al., 2017
  - Bradley et al., 2016
  - Suarez et al., 2013
  - Tzankova & Schiphorst, 2010
  - Forlini et al., 2016
Overview (cont.)

- Research methods in digital humanities (http://www.library.illinois.edu/sc/services/Digital_Humanities/Research_Methods.html)
  - Data curation
  - Digital publishing
  - Image analysis
  - Text encoding
  - Machine learning
  - **Data visualization**
  - Geospatial information systems
  - Text mining
  - ...

Overview (cont.)

• Data/Information Visualization

  • The technique of creating “2- or 3-dimensional representations of data that enable new discoveries of insights and knowledge” (Soukup & Davidson, 2002)

  • Unique capabilities for...
    • Pattern recognition
    • Outlier detection
    • Knowledge/insight discovery
Overview (cont.)

• Information Visualization and Digital Humanities
  
  • Information visualization can be used to analyze metadata, tags, texts, images, or any type of digital information in humanities works
  
  • Information visualization has been considered as one of the core skills for librarians to support research in various disciplines
  
  • How about using information visualization in digital humanities research?
Overview (cont.)

• Research questions

  • Q1: What are the main tools, application types, and purposes of information visualization in digital humanities research?

  • Q2: What types of online resources are provided by academic libraries that focus on applying information visualization to digital humanities research?

  • Q3: Is there a gap between the practical use of information visualization in digital humanities research and the support and services provided by academic librarians? And if so, how can academic librarians prepare best?
Literature Review

- Information visualization used in the digital humanities
  - screenplay parsing (Hoyt et al. 2014)
  - religion symbol analysis (Tzankova & Schiphorst 2010)
  - history understanding (Forlini et al. 2016; Suarez et al. 2013)
  - web-information visualization (Breure et al. 2014; Lynch 2014)

- Information visualization discussed in the LIS community
  - a core skill and focus area to be incorporated into information literacy (Womack 2014; Roff 2011)
  - a tool for teaching research skills (Huff 2007)
  - a practice to create a meaningful visual storytelling experience (Magnuson 2016)

- The research gap...?
Methodology

- Content analysis of...
  - five years of recognized journals in DH
  - online resources provided by academic libraries in support of DH research and practices

- Compare tools, applications types, and purpose of IV use in DH research
Data: DH Journals

- Selecting DH journals
  - peer-reviewed, reached a wide audience, and fully accessible through my university library subscriptions

- Using supporting resources
  - Resource Guide provided by Digital Humanities at Berkeley, The Digital Humanities: Journals provided by USC Libraries Research Guides, and Digital Humanities: Journals provided by UIC University Library Research and Subject Guides

- Cross-comparing the lists
  - Three journals were selected: Digital Humanities Quarterly (DHQ), Digital Scholarship in the Humanities (DSH) (formerly known as Literary and Linguistic Computing), and Digital Studies / Le champ numérique (DS/CN)

- Identifying IV-related articles
  - whether the article’s title, abstract, or keyword mentioned information visualization, data visualization, visual analytics, visualization, or visualizing
  - whether information visualization techniques were used to address the article’s research needs.
Data: DH Journals (cont.)

<table>
<thead>
<tr>
<th>Journal Name</th>
<th>Format/Type</th>
<th># of Issues per Year (2012-2016)</th>
<th>Total # of Published Articles (2012-2016)</th>
<th>Total # of InfoVis-related Articles (2012-2016)</th>
<th>% of InfoVis-related Articles (2012-2016)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DHQ</td>
<td>Open access, peer-reviewed, online only</td>
<td>3-4</td>
<td>138</td>
<td>49</td>
<td>26.2%</td>
</tr>
<tr>
<td>DSH</td>
<td>Peer-reviewed, print &amp; online</td>
<td>4</td>
<td>205</td>
<td>72</td>
<td>25.99%</td>
</tr>
<tr>
<td>DS/CN</td>
<td>Open access, peer-reviewed, online only</td>
<td>1-3 (rolling basis)</td>
<td>53</td>
<td>17</td>
<td>24.29%</td>
</tr>
</tbody>
</table>
Data: DH Online Resources

• According to Adams and Gunn (2013),
  
  • “DH invites—and demands—collaboration with parties outside of the library”
  • “Librarians need additional training and education in order to contribute effectively”

• Academic library websites with a focus on both DH and IV were considered

• Three university library guides were selected for analysis:
  
  USC Libraries Research Guides for The Digital Humanities: Digital Visualization
  University of Tennessee-Knoxville University Libraries Research Guides for Digital Humanities: Visualizations
  Cornell University Library Research Guides for Digital Humanities: Data Visualization
Results: DH Journals

• Three categories of IV use were examined: tools, applications, and purposes

• IV Tools
  
  • At an aggregate level, approximately 70% of the visualization tools used in DH research were free and open-source
    • E.g., Neatline for geo-temporal interpretation of archival collections (Nowviskie et al. 2013) and ANNIS for visualizing complex multi-layer linguistic corpora (Krause and Zeldes 2016)
    • More than half was designed and developed by the author(s)

  • About 30% of the visualization tools used enterprise software
    • E.g., Carrot for visualizing and making sense of a large Web collection through cluster analysis (Miligan et al. 2016)
Results: DH Journals (cont.)

• IV Applications (broadly defined)

• For example,

  • “document analysis” application includes analysis of not only textual data such as corpus and archives but also screenplays and hypertexts.

  • “history” applications cover studies ranging from literary history to art history
Results: DH Journals (cont.)

- IV Applications (broadly defined)
Results: DH Journals (cont.)

• IV Purposes

• Based on Agrawala (2005), there are three main purposes of information visualization use

  • presenting data in context: communicate existing information to others
  
  • analysis and discovery: using visualization techniques to help discover new knowledge, insights, and even questions
  
  • storytelling: designing a graphic that connects to and resonates with users through an overarching theme
Results: DH Journals (cont.)

• IV Purposes
## Results: University Library Research Guides

<table>
<thead>
<tr>
<th>Information visualization components</th>
<th>USC Libraries Research Guides for The Digital Humanities: Digital Visualization</th>
<th>University of Tennessee-Knoxville University Libraries Research Guides for Digital Humanities: Visualizations</th>
<th>Cornell University Library Research Guides for Digital Humanities: Data Visualization</th>
</tr>
</thead>
<tbody>
<tr>
<td>Definition</td>
<td>✔</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Introduction</td>
<td>✔</td>
<td>✔</td>
<td></td>
</tr>
<tr>
<td>Journals</td>
<td>✔</td>
<td>✔</td>
<td></td>
</tr>
<tr>
<td>Apps/Tools</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>Data resources</td>
<td>✔</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Related books</td>
<td></td>
<td></td>
<td>✔</td>
</tr>
<tr>
<td>Coding and software resources</td>
<td></td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>Human resources</td>
<td></td>
<td></td>
<td>✔</td>
</tr>
</tbody>
</table>
Results: University Library Research Guides (cont.)
Insights and Discussions

• The “depth” of IV use matters more—namely how IV has been applied (applications), and why it has been adopted (purposes).

• It is important to recognize the wide variety of applications for IV in DH research.
  • The application of IV is strongly correlated with the humanities data under consideration, whether it be corpus, archives, music, or artwork.

• For most of IV-related work in DH, the purpose is to analyze and discover.
  • What is the most accurate way to interpret the visualization results?
  • What can be deduced or mined from these results?
  • How can new discoveries and insights be drawn from the visualizations?
Insights and Discussions (cont.)

• Be familiar with common data types in digital humanities research, and understand what kinds of data work best with each visualization technique

• Capitalize on online courses such as MOOCs to hone the skills in using visualization apps and tools

• Recognize the importance of applying analytical thinking to make the best possible interpretation of the visualization results

• Never be afraid to reach out for the “human resources” (i.e., the domain experts) for collaboration opportunities
• **Voyant** ([http://voyant-tools.org/](http://voyant-tools.org/)) for text analysis and visualization
References

Q&A

Thank you in various languages:
- danke
- 謝謝
- ngiyabonga
- teşekkür ederim
- gracias
- grazie
- go raibh maith agat
- merci
- obrigado
- terima kasih
- 감사합니다
- धन्यवाद
- شكراً جزيلاً