Building Data Collections for Digital Humanities Research

Devin Higgins
MSU Libraries
higgi135@msu.edu
@devinhhi
Google Ngrams: “Humanities Data”
A wide plain, where the broadening Floss hurries on between its green banks to the sea, and the loving tide, rushing to meet it, checks its passage with an impetuous embrace. On this mighty tide the black ships—laden with the fresh-scented fir-planks, with rounded sacks of oil-bearing seed, or with the dark glitter of coal—are borne along to the town of St. Ogg's, which shows its aged, fluted red roofs and the broad gables of its wharves between the low wooded hill and the river brink, tinging the water with a soft purple hue under the transient glance of this February sun. Far away on each hand stretch the rich pastures, and the patches of dark earth, made ready for the seed of broad-leaved green crops, or touched already with the tint of the
What can topic models of PMLA teach us about the history of literary scholarship?
Andrew Goldstone and Ted Underwood
Frankfurt School: Minima Moralia by Theodor Adorno 1951

Theodor Adorno

Minima Moralia: Dedication

The melancholy science, from which I make this offering to my friend, relates to a realm which has counted, since time immemorial, as the authentic one of philosophy, but which has, since its transformation into method, fallen prey to intellectual disrespect, senseless caprice and in the end forgetfulness: the teaching of the good life. What philosophy once called life, has turned into the sphere of the private and then merely of consumption, which is dragged along as an addendum of the material production-process, without autonomy and without its own substance.

http://voyant-tools.org/
"non-consumptive research" is text mining or computational "analysis...on one or more Books, but not research in which a researcher reads or displays substantial portions of a Book to understand the intellectual content presented within the Book."

*Google v. the Authors Guild and the Association of American Publishers Settlement Agreement, 2009*
After 10 Years, Google Books Is Legal

Thanks to a landmark ruling, information just got a little more free.

ROBINSON MEYER | OCT 20, 2015
THE SILVER LINING.


Index to this number.


No sooner was the memorial presented to Congress than opposition to the movement began to manifest itself among railroad officials and attorneys, who filed suit in Washington to look after the interests of their roads and their clients. Merchants, and other business men whose success depends upon the prosperity of agriculture, are beginning to look to this question square in the face, and with few exceptions strike hands with the 'Grangers,' and are enrolling themselves with the thousands that are now petitioning Congress for national action and control of inter-state commerce. There are some of the leading political papers of all the different political parties now actively advocating the measure, and Boards of Trade are not idle lookers-on in this most important movement ever inaugurated in the interest of the producers and consumers in this country: but we know men as they are, they are looking to their own interests, and not specially to the interests of farmers or other business men whose influence is necessarily interconnected with agriculture.

The New York Board of Trade and Transportation has appointed a committee to investigate this grand movement and examine the Grange petitions and bills now pending before Congress, and recommend such legislation as the "Board" may deem to be necessary.

The Transportation Question.

E. J. Woodman, Paw Paw.

It is well known to every Patron of the "Board" that the question is being given serious attention by that body. The "Board" is now a public body and is appointed to consider and act upon the transportation question, and the Grange has been requested to present its views and recommendations to the "Board." The Grange has accordingly sent a delegation to the "Board," and the delegation has presented its views and recommendations to the "Board." The "Board" has received the views and recommendations of the Grange, and has promised to consider them carefully and act upon them.

No sooner was the memorial presented to Congress than opposition to the movement began to manifest itself among railroad officials and attorneys, who filed suit in Washington to look after the interests of their roads and their clients. Merchants, and other business men whose success depends upon the prosperity of agriculture, are beginning to look to this question square in the face, and with few exceptions strike hands with the 'Grangers,' and are enrolling themselves with the thousands that are now petitioning Congress for national action and control of inter-state commerce. There are some of the leading political papers of all the different political parties now actively advocating the measure, and Boards of Trade are not idle lookers-on in this most important movement ever inaugurated in the interest of the producers and consumers in this country: but we know men as they are, they are looking to their own interests, and not specially to the interests of farmers or other business men whose influence is necessarily interconnected with agriculture.

The New York Board of Trade and Transportation has appointed a committee to investigate this grand movement and examine the Grange petitions and bills now pending before Congress, and recommend such legislation as the "Board" may deem to be necessary.

The Transportation Question.

E. J. Woodman, Paw Paw.

It is well known to every Patron of the "Board" that the question is being given serious attention by that body. The "Board" is now a public body and is appointed to consider and act upon the transportation question, and the Grange has been requested to present its views and recommendations to the "Board." The Grange has accordingly sent a delegation to the "Board," and the delegation has presented its views and recommendations to the "Board." The "Board" has received the views and recommendations of the Grange, and has promised to consider them carefully and act upon them.
Humanities Data

MSU Libraries Humanities data includes but is not limited to digitized and born digital text, audio, images, moving images, and the metadata that describes them. Current collection strengths reside in text and audio data. The following collections have been prepared with an eye toward enabling computational analysis at the micro and macro scale.

Recommendations for acquisition of new Humanities data, requests for assistance gathering and preparing Humanities data, and questions about how to use Humanities data should be directed to the MSU Libraries Digital Humanities team.

Text

- Sunday School Books in Nineteenth Century America
  - Number of Works: 166 works
  - Years covered: 1809-1887
  - Size: 11.6 MB

- The Grange Visitor
  - Number of Works: 429 issues
  - Years covered: 1875-1896
  - Size: 8.53 GB

[lib.msu.edu/dh/humdata](lib.msu.edu/dh/humdata)
Thomson Reuters Web of Science
“Lite” (free) API to return metadata,
“Premium” (costly) API to return citation data.

JSTOR
Web-Accessible word-frequencies and metadata, full
text by request/contract

Gale
Data available for purchase on “cost-recuperative”
basis

APIs and Bulk Text Access
<table>
<thead>
<tr>
<th>Timestamp</th>
<th>Series</th>
<th>Collection</th>
<th>(text) Collection Size</th>
<th>Indicate scale</th>
<th>(image) Collection size</th>
<th>Indicate scale</th>
<th>Collection Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>9/29/2015 12:54:30</td>
<td>Sabin Americana: 1500-1926 (backfile through 2010)</td>
<td>147 GB</td>
<td>1673 GB</td>
<td></td>
<td></td>
<td></td>
<td>Based on Joseph Sabin's famed bibliography. Bibliotheca Americana. Covering a span of 400 years in North, Central and South America. This digital collection consists of testimony and documents. The handwritten documents within Personal Justice Denied, Public Hearings of the Committee are particularly valuable. Part IV: The Age of Emancipation, the fourth installment. In its entirety, Slavery and Anti-Slavery will consist of more than 500 pages of text and images. This collection allows you to explore the internal organization. Users have access to the complete run of The Times Literary Supplement until 1974, writings and influential criticism of hundreds of books. The value of TLS Historical Archive lies in its extensive collection of literature. This unique digital collection offers thousands of books and periodicals.</td>
</tr>
<tr>
<td>9/29/2015 13:02:22</td>
<td>Archives Unbound</td>
<td>Personal Justice Denied: Public Hearings of the Committee</td>
<td>51 MB</td>
<td>1.8 GB</td>
<td></td>
<td></td>
<td>Date Range: 1966-1978 Content: 26,577 pages Source Library: Amistad Research Center Fannie Lou Hamer was an voting rights activist and civil rights leader who fought against segregation, disfranchisement, and violence in the Deep South. The collection includes a wealth of documents and images related to her work. The largest and most ambitious project of its kind, Slavery and Anti-Slavery provides a comprehensive overview of the legal, economic, and social aspects of slavery in the Atlantic World. Part II: The Slave Trade in the Atlantic World continues to expand with new acquisitions. Date Range: 1960-1986 Content: 10,000 pages Source Library: Auburn Avenue Research Library, Atlanta Description: Amin Baraka is the author of over 40 books of essays, poems, and fiction. His works explore the complexities of race, identity, and politics in the United States and around the world.</td>
</tr>
</tbody>
</table>
Thanks!

Building Data Collections for Digital Humanities Research