2 April, 11:00am

Digital humanities

Dr Miguel Escobar Varela
Digital Humanities
Challenges for libraries

Miguel Escobar Varela, Ph.D.
Assistant Professor, National University of Singapore
Academic Advisor on Digital Scholarship, NUS Libraries
miguelescobar.com
@miguelJogja
This talk

What are the digital humanities?

What this means for libraries: challenges and opportunities

Some ideas for how to tackle these challenges and capitalize on these opportunities
What are the digital humanities (DH)?

The usage of digital data and digital methods for carrying out research in the humanities
Areas in the DH

Databases

Data-driven analysis

Digital publishing

The DH are also a community, a space for meeting up and discussing how technology is changing the humanities
Databases
Databases

Data is ‘potential information’. A process needs to be applied to data to become meaningful. This is a challenge to how people usually think about research in the humanities.
Data is context dependant

- Data is context-dependant
- The question should not be what are data but "when are data" (Borgman, 2015)

That which is data in one context might be information or knowledge in another.
Sources of data in the humanities

- Digitized books, films, artworks, video recordings
- Motion capture
- 3d models
- Surveys
- Production details (film, theatre)
- Historical records
Problems of DH data

Standardization

Incompleteness

Inaccuracy

These problems are important for historical data.

Annotation and discussion are the best way to deal with this problems.
Limited data in the humanities

20-30,000 novels published

6,000 novels still extant

3,500 novels available in full text

Digitization and availability of British Nineteenth-Century Novels. Image published under a CC-BY license (Schöch).
The future of data in DH

From Schöch (2013)
Smart data

Markup, annotations and metadata

Clear data models

Clear relations to external entities

“Smart data to be semi-structured or structured, clean and explicit, as well as relatively small in volume and of limited heterogeneity” Schöch
The Comédie-Française Registers Project

[Image of the Comédie-Française website]

---

<table>
<thead>
<tr>
<th>Titre</th>
<th>Auteur</th>
<th>Genre</th>
<th>Nombre de représentations</th>
<th>Nombre de reprises</th>
<th>date de la première</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tartuffe ou l'Imposteur</td>
<td>Jean-Baptiste Poquelin dit Molière</td>
<td>comédie</td>
<td>956</td>
<td>0</td>
<td>1663-04-03</td>
</tr>
<tr>
<td>Médécin malgré lui (Le) / Le Médecin forcé (titre alt.)</td>
<td>Jean-Baptiste Poquelin dit Molière</td>
<td>comédie</td>
<td>692</td>
<td>0</td>
<td>1669-10-31</td>
</tr>
<tr>
<td>École des femmes (L')</td>
<td>Jean-Baptiste Poquelin dit Molière</td>
<td>comédie</td>
<td>744</td>
<td>0</td>
<td>1665-06-25</td>
</tr>
<tr>
<td>Crispin médecin</td>
<td>Hectarache (Noël Lebrun, sieur de)</td>
<td>comédie</td>
<td>735</td>
<td>0</td>
<td>1669-12-12</td>
</tr>
<tr>
<td>Espirito de contracion (L')</td>
<td>Du Fossey (Charles)</td>
<td>comédie</td>
<td>690</td>
<td>0</td>
<td>1700-08-27</td>
</tr>
<tr>
<td>Avocat Pataquin (L')</td>
<td>Bruyré (David-Augustin de)</td>
<td>comédie</td>
<td>685</td>
<td>0</td>
<td>1706-06-04</td>
</tr>
<tr>
<td>Asane (L')</td>
<td>Jean-Baptiste Poquelin dit Molière</td>
<td>comédie</td>
<td>674</td>
<td>0</td>
<td>1700-05-14</td>
</tr>
<tr>
<td>George Dandin ou le Mai confondu</td>
<td>Jean-Baptiste Poquelin dit Molière</td>
<td>comédie</td>
<td>682</td>
<td>0</td>
<td>1800-05-15</td>
</tr>
<tr>
<td>Plaidiers (Les)</td>
<td>Jean Racine</td>
<td>comédie</td>
<td>645</td>
<td>0</td>
<td>1890-09-04</td>
</tr>
<tr>
<td>École des maris (L')</td>
<td>Jean-Baptiste Poquelin dit Molière</td>
<td>comédie</td>
<td>636</td>
<td>0</td>
<td>1990-05-16</td>
</tr>
<tr>
<td>Joueur (Le)</td>
<td>Regnard (Jean-François)</td>
<td>comédie</td>
<td>633</td>
<td>0</td>
<td>1990-12-19</td>
</tr>
<tr>
<td>Amphitryon</td>
<td>Jean-Baptiste Poquelin dit Molière</td>
<td>comédie</td>
<td>636</td>
<td>0</td>
<td>1860-04-06</td>
</tr>
<tr>
<td>Misanthrope (Le)</td>
<td>Jean-Baptiste Poquelin dit Molière</td>
<td>comédie</td>
<td>591</td>
<td>1</td>
<td>1990-05-27</td>
</tr>
<tr>
<td>Florestin (Le)</td>
<td>La Fontaine (Jean de)</td>
<td>comédie</td>
<td>564</td>
<td>0</td>
<td>1990-07-23</td>
</tr>
<tr>
<td>Comtesse d'Esclarones (La)</td>
<td>Jean-Baptiste Poquelin dit Molière</td>
<td>comédie</td>
<td>548</td>
<td>0</td>
<td>1990-05-25</td>
</tr>
</tbody>
</table>

https://www.cfregistes.rs.org/en/the-data/basic-tool
Data-driven research

Network analysis, geospatial visualizations, time series analysis and textual analysis
Networks

A network consists of:

Nodes (things that are connected).

Edges (connections between those things). Specific, explicit connections between things. They can be directed or undirected.

Examples: social networks, communication networks, citation networks, collaboration networks.
Network measurements

**Degree.** Total number of edges it has to other nodes.

**Density.** The portion of the potential connections in a network that are actual connections.

**Average path length.** The average steps (edges) to go from one node to another node.
Network analysis of wayang kulit characters

Network analysis of Javanese wayang kulit characters

<table>
<thead>
<tr>
<th></th>
<th>Percentage</th>
<th>Average Weighted Degree</th>
</tr>
</thead>
<tbody>
<tr>
<td>Indian Characters</td>
<td>53%</td>
<td>153.47</td>
</tr>
<tr>
<td>Javanese Characters</td>
<td>47%</td>
<td>73.36</td>
</tr>
</tbody>
</table>
Textual analytics

Many different approaches (most developed area of DH).

- Corpus linguistics
- Topic modelling
- Sentiment analysis
Voyant Tools (voyant-tools.org)

Analysis of letters by Vincent Van Gogh (1853-1890)

https://voyant-tools.org/?corpus=2a9aa299a95d7eca47cf68d25f0382e7
Robots reading vogue

Bookworm
http://bookworm.library.yale.edu/

Topic modelling
http://dh.library.yale.edu/projects/vogue/topics/
Geospatial analysis

Increasingly common approach

- Geospatial visualizations
- Geostatistics
Visualization of wayang kulit performances (Java, Indonesia)
Spatial distribution of Chinese Culture in Singapore

http://shgis.nus.edu.sg/
Time-series analysis

https://www.cfregisters.org/en/the-data/basic-tool
Javanese Wayang Kulit Performances Per Year
Digital publishing
Titus Andronicus, Hong Kong Arts Festival and No Man's Land
Director: TANG Shu-wing, Date: 2008

Pathfinders (project built with Scalar)

http://scalar.usc.edu/works/pathfinders/traversals-and-interviews-documentation
The Chinese Deathscape

Grave Reform in Modern China

In the past decade alone, more than ten million corpses have been exhumed and reburied across the Chinese landscape. The campaign has transformed China’s graveyards into sites of acute personal, social, political, and economic contestation.

In this digital volume, three historians of China, Jeffrey Snyder-Flanholt, Christian Henrick, and Thomas S. Mullaney, chart the history of China’s rapidly shifting deathscape. Each essay grapples with a different dimension of grave relocation and burial reform in China over the past three centuries: from the phenomenon of “baby towers” in the Lower Yangzi region of late imperial China, to the histories of death in the city of Shanghai, and finally to the history of grave relocation during the contemporary period, examined by Mullaney, when both its scale and...
Challenges and suggestions
Challenges for libraries

Changing roles for people and spaces

Sustainability

Development support
Two ideas for discussion

- The library should be the home of DH in the university
- A simple distinction between researcher and librarian doesn’t make sense in the 21st century
Some suggestions for libraries

Rethink the distinction between librarians and academics

Rethink the library as a space for dialogue and exchange
Thank you

Comments and questions

m.escobar@nus.edu.sg

miguelescobar.com

@miguelJogja