Mapping Knowledge Transmission in Imperial China: Data Development and Data Integration in Chinese History

Hilde De Weerdt
University of Oxford
I. Research methods: traditional and digital historical research
Locations of notebook printing in Southern Song

Places of origin of interlocutors in *Huizhu lu*
Timeline of notebook printing, 1150-1250

Datestamp of entries in *Huizhu lu 2*
Huizhu lu 2 quotations by type and subtopic

Song printed notebooks by century of completion and century of printing
Network of authors and interlocutors by citation frequency in *Huizhu lu* 1 and 2.

Network of producers of printed Song *biji*
### Comparison of character frequencies in different editions

<table>
<thead>
<tr>
<th>Term</th>
<th>Siku freq</th>
<th>Zhonghua freq</th>
</tr>
</thead>
<tbody>
<tr>
<td>金人</td>
<td>47</td>
<td>30</td>
</tr>
<tr>
<td>大金</td>
<td>22</td>
<td>22</td>
</tr>
<tr>
<td>金國</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>虜</td>
<td>2</td>
<td>157</td>
</tr>
<tr>
<td>女真</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>契丹</td>
<td>14</td>
<td>13</td>
</tr>
<tr>
<td>夷</td>
<td>18 (3)</td>
<td>34</td>
</tr>
<tr>
<td>狄</td>
<td>2</td>
<td>12</td>
</tr>
<tr>
<td>蠻</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>戎</td>
<td>16</td>
<td>21</td>
</tr>
<tr>
<td>醜</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>盟</td>
<td>12</td>
<td>16</td>
</tr>
<tr>
<td>酋</td>
<td>1</td>
<td>10</td>
</tr>
<tr>
<td>敵人</td>
<td>49</td>
<td>2</td>
</tr>
</tbody>
</table>

### Source text type

<table>
<thead>
<tr>
<th>Source text type</th>
<th>Number of occurrences</th>
<th>Number of entries</th>
</tr>
</thead>
<tbody>
<tr>
<td>War announcement</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>Military institutions survey</td>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td>Pro-war memorial</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Subtotal</td>
<td>16</td>
<td>3</td>
</tr>
<tr>
<td>Other (memorial, envoy report, poem, memoir, anecdote, conversation, notebook entry copy)</td>
<td>20</td>
<td>15</td>
</tr>
</tbody>
</table>

### Collocates

<table>
<thead>
<tr>
<th>freq</th>
<th>score</th>
<th>word</th>
</tr>
</thead>
<tbody>
<tr>
<td>30</td>
<td>16.2</td>
<td>人</td>
</tr>
<tr>
<td>22</td>
<td>16.9</td>
<td>大</td>
</tr>
<tr>
<td>5</td>
<td>4.4</td>
<td>國</td>
</tr>
<tr>
<td>4</td>
<td>14.5</td>
<td>寇</td>
</tr>
<tr>
<td>2</td>
<td>3.6</td>
<td>贼</td>
</tr>
<tr>
<td>2</td>
<td>1.2</td>
<td>使</td>
</tr>
<tr>
<td>1</td>
<td>0.5</td>
<td>兵</td>
</tr>
</tbody>
</table>

**LINGUISTIC ANALYSIS**
Digital text analysis

XML

Text analysis software
II. Prosopographical data and the analysis of notebook producers
A. OFFICIAL CAREER
The number of court postings of Northern Song printed *biji* authors by type of posting.
The number of court postings of Southern Song printed *biji* authors by type of posting.
B. ENTRY
Percentage of methods of entry into officialdom for Song *biji* authors
(46 out of 73 unique authors)
Number of *biji* authors by type of entry into officialdom shown by date of entry.

Note: dates of entry available for 40 out of the 46 cases in *CBDB*.
Numbers of *biji* authors printed in Song times for whom no entry data exist in *CBDB*.
C. CBDB RELATIONSHIPS
Number and type of associations for all producers of printed Song *biji*. 
Network of printed Song *biji* producers showing tie strength.
Network of printed Song *biji* producers (excluding Yang Wanli) by clusters.
Size of nodes marks degree of Eigenvector centrality.
• Problems
  coverage of biographical and relationship data

• Hypotheses
  author backgrounds
  network development
Wang Mingqing’s (1127- post 1214) relationships to the first degree in CBDB (Pajek GBK)
Wang Mingqing’s relationships to the first degree in CBDB (exported to ARCGIS)
II. Text encoding and the analysis of information networks in notebooks
A. The Text as Database
[Chinese text]
Authors in red; interlocutors in green. Outer circle: 1 reference; inner circles: 2 (light green) to 3 (dark green) and so up to 15 references.
Author and Interlocutor Network
(Huizhu 1 and 2) Pajek
Author network (authors to 959)
Author network (authors 960-1059)
Author network (authors 1060-1159)
Author network (authors 1100-1199)
Interlocutors’ residential addresses

Dark brown areas mark houlu interlocutors, white areas qianlu interlocutors, striped areas include both qianlu and houlu interlocutors. Numbers indicate number of resident interlocutors (multiple residencies included).
Interlocutors’ office addresses

Dark brown areas mark county and prefectural postings; shaded areas mark regional postings. Dot sizes mark number of postings (1-3).
Interlocutors: local vs. regional offices (houlu)
Secondary interlocutors:
office types by jurisdiction (houlu)
• Problems
  time
  place
  office

• Some preliminary results
  size and composition of oral information network
  composition of citation network of written sources

• Hypothesis
  horizontal dimensions of imperial information networks
III. Conclusions and Future Prospects
Project overview

Bibliographic and citation analysis

- Notebooks in print

Analyze by

- Time: compilation and printing
- Subject: primary through tertiary
- Printer: type
- Place: prefecture
- Combinations of these

Search by

- Author
- Title
- Cited titles: name and genre
- Citations: topic
- Citations: type
- Combinations of these

Notebooks in xml: citations and quotations

(time) (table)
(table)
(table)
(map) (table)
(map) (timeline) (table)
(results list)
(results list)
(table)
(table)
(table)
Network analysis

- **Notebooks in print**
  - (table with author and editor IDs to allow further SNA through CBDB)

- **Notebooks in xml: citations and quotations**
  - Analyze by
    - Person(s): type(s) and name(s): authors, interlocutors, collectors citing and cited
      - (network diagram) (table)
    - Topic(s): authors, interlocutors, and collectors citing or cited
      - (network diagram) (table)
    - Time: fl. date of authors, interlocutors, and collectors citing or cited
      - (network diagram) (table)
    - Place: residential and office addresses of authors, interlocutors, collectors citing and cited
      - (map) (table)
    - Combinations of these
      - (map) (network diagram) (table)
• Advantages of linking large datasets with small-scale projects

• Focus on spatial, network, numeric analysis rather than digital publishing

• Benefits of multidimensional and multiperspectival analyses
References and Credits
