Chinese Internet: A Cultural Sphere, National Sphere, or Transnational Public Sphere

by Han-Teng Liao

For City University of Hong Kong
Curriculum Vitae

Han-Teng Liao
Oxford Internet Institute
Oxford OX1 3JS, UK
EMAIL: hanteng@gmail.com
WEB: http://people.ox.ox.ac.uk/hanteng
MOBILE: +44 7467266819
SKYPE: hanteng.liao.ouo
LINKEDIN: Han-Teng Liao

EDUCATION

University of Oxford, Oxford, UK
DPhil (PhD) Information, Communication and the Social Sciences, July 2014
Thesis supervisor: Ralph Schroeder
Thesis examiners: William H. Dutton and Mike Thelwall

National Taiwan University, Taipei, Taiwan
M.A. Journalism, (GPA 4.0), June 2004
M.Sc. Computer Science and Information Engineering (GPA 3.9), June 2004
B.A. Foreign Languages and Literatures, June 2000
B.Sc. Electrical and Electronic Engineering, June 2000
Expertise

- Webometrics; User-generated content; Wikipedia research; Information Science in China and Taiwan; Chinese Internet; Chinese Web data;
- Open Data; Open Source; Open Knowledge; Software, Website and content internationalization-localization (i18n/L10n); Digital Chinese characters
- Geo-linguistic regions and analysis
Research agenda:

How can (Chinese) Internet conceptualized as a (Chinese) cultural sphere, national sphere, or transnational public sphere?

- Guobin Yang’s (2003) notion of a Chinese cultural sphere
- e.g. National web studies, including “Mapping Iran Online” and “Big UK Domain Data”
- Transnational public sphere: “A growing body of media studies literature is documenting the existence of discursive arenas that overflow the bounds of both nations and states.” (Fraser, 2007)
- A survey: http://goo.gl/vcyngX delineating and defining Chinese Web
Tentative hypothesis:

Beijing’s authoritarian persistence is positively validated by the “boundary work” of national sphere within mainland China

The diversity of (transnational) Chinese cultural sphere, especially elements contributed by Hong Kong and Taiwan remains essential for the possibility of a (transnational) public sphere.
What's gonna bring down the totalitarian government in China? Is it gonna be guns? Is it gonna be tanks? Is it gonna be missiles? No it's gonna be KJ. KJ is a 25-year-old girl who lives in Taiwan, ... She is talking to the mainland Wikipedians, and she's helping them to get through the firewalls. .... Every day with something she is doing, and every day she is bringing in a bit of democracy and information of freedom to China. … This is going on all the time in hundreds and thousands of ways all over the world. All of these censorship regimes are porous. They are becoming more porous and they will be more porous [sic].
## Four Major Chinese-speaking Regions

<table>
<thead>
<tr>
<th>Country Code</th>
<th>Political Filtering(^a)</th>
<th>Networked Readiness Index Ranking(^b)</th>
<th>Democracy World Ranking(^c)</th>
<th>Free Speech World Ranking(^d)</th>
<th>Major Human Rights Violation(^e)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CN (Mainland China)</td>
<td>Pervasive</td>
<td>59</td>
<td>138</td>
<td>163</td>
<td>Yes</td>
</tr>
<tr>
<td>HK (Hong Kong)</td>
<td>None</td>
<td>22</td>
<td>78</td>
<td>61</td>
<td>No</td>
</tr>
<tr>
<td>SG (Singapore)</td>
<td>None</td>
<td>5</td>
<td>84</td>
<td>141</td>
<td>No</td>
</tr>
<tr>
<td>TW (Taiwan)</td>
<td>None</td>
<td>13</td>
<td>32</td>
<td>32</td>
<td>No</td>
</tr>
</tbody>
</table>

\(^a\) Data from the OpenNet Initiative about Internet filtering of political contents.

\(^b\) Data from the Networked Readiness Index 2006-2007 Ranking, by World Economic Forum, among 122 countries surveyed.

\(^c\) Data from the Economist Intelligence Unit's index of democracy 2006, with 167 countries are ranked from the top democracies to authoritarian regimes.

\(^d\) Data from the Worldwide Press Freedom Index 2007 made by the Reporters without Borders, with 169 countries ranked.

\(^e\) Whether a state is mentioned in the Human Rights Watch World Report 2007.
“Modern China: A Very Short Introduction”

- China itself is neither fully free nor democratic. Taiwan, since the 1990s, has been both free and democratic. Singapore, a largely Chinese society, is democratic, in that it has regular elections which are nominally open to opposition candidates (but at high cost to themselves), but is not free (the media and political activism are both heavily regulated). Most intriguing is Hong Kong, which is little more democratic than it was under the British. Yet it is a very free society: although there is political pressure and a certain level of self-censorship, it has a lively press, it is easy to publish books attacking the Chinese government, and it supports a variety of political parties (although the legislature is arranged to prevent any such party ever coming to power). There are few, if any, other such free, undemocratic societies.
CULTURAL POLITICS OF
USER-GENERATED ENCYCLOPAEDIAS:
Comparing Chinese Wikipedia and Baidu Baike

by Han-Teng Liao

How have the two major Chinese-written user-generated encyclopaedias, Baidu Baike and Chinese Wikipedia, overcome, reinforced or shifted the existing cultural-political boundaries among mainland China, Hong Kong and Taiwan?
How has the Web overcome, reinforced or shifted the existing cultural-political boundaries?

General Question
How have *Baidu Baike* and *Chinese Wikipedia* overcome, reinforced or shifted the existing cultural-political boundaries among China, Hong Kong and Taiwan?

Research Question
Findings: some highlights

Chinese Wikipedia:
- integration patterns and designs (with ongoing tensions)
- boundary-overcoming patterns

Baidu Baike:
- culturally-and-politically sanitized domestic market
- boundary-reinforcing patterns
- Web spheres,
- Processability
- Cultural thickening

- Editorial
- Content/Citation
- Reception/Use

- The impact of the Great Firewall
- Boundaries online
Structuring the web

- Noortje Marres (2009): “web services incorporate social science methods like textual analysis, social network analysis, and geospatial analysis, arguably ordering data for … research.”

- Richard Rogers (2013, p.207):
  - May the web deliver structured data after all?
  - In this way of thinking, web services -search engines, collaboratively authored wikis, and social networking platforms- become the data filterers, leaning and ordering the data for end use as well as perhaps for research.
## Web spheres

<table>
<thead>
<tr>
<th>Language only</th>
<th>Region only</th>
<th>Platform only</th>
<th>Language and platform</th>
<th>Region and platform</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spanish-speaking</td>
<td>U.S. cyber domain</td>
<td>Wikipedia</td>
<td>Chinese Wikipedia</td>
<td>Encyclopedia of Taiwan</td>
</tr>
<tr>
<td>web-sphere</td>
<td></td>
<td></td>
<td>Spanish Wikipedia</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Arabic Wikipedia</td>
<td>The Encyclopedia of Hong Kong</td>
</tr>
<tr>
<td>Chinese-language</td>
<td>China's cyberspace</td>
<td>Baidu Search</td>
<td>Chinese Weibo</td>
<td>Hong Kong Virtual Communities</td>
</tr>
<tr>
<td>Internet</td>
<td></td>
<td>Yandex</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Google Search</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Google News</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Google Books</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arabic online world</td>
<td></td>
<td>Twitter universe</td>
<td>Arabic Twitter</td>
<td>Hong Kong Weibo</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Facebook nation</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Craigslist</td>
<td></td>
<td>Craigslist Hong Kong</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Web spheres: specify the research object

Virtual methods (Schneider & Foot, 2005)
Digital methods (Rogers, 2013)

Major Chinese geolinguistic regions

- zh-cn: mainland China
  - media system 1 (nation) state 1

- zh-hk: Hong Kong
  - media system 2 (nation) state 2

- zh-tw: Taiwan
  - media system 3 (nation) state 3
Web spheres
(Rogers, 2013)

- “Web sphere” analysis, initially, did not refer to search engine space, but rather a meticulous collection of thematically related sites for further analysis (Rogers, 2013, p52).

- As scholars continue to disaggregate the online (as search engines already have done in providing separate subengines by sphere), the concerns shift away from the study of Internet and politics in general to the politics of separate spaces (Rogers, 2013, p59).
Processability: combined human- and computer-manipulation

Figure 3-1. Processability shapes web spheres
Increasing processability
Types of texts and technologies: a comparison
East Asians are accustomed to dealing with a multibyte system, in contrast to Western monobyte reductionist culture. It may be that in the future our multibyte culture will prove advantageous for dealing with complex systems.

- East Asian Science Technology and Society (STS) scholar, Nakayama Shigeru
Figure 3-2. Cultural thickening patterns that overcome or reinforce boundaries

Cultural thickening: intensified patterns of ties

(Couldry & Hepp, 2012)

Definitions

Findings

Conclusions

Implications

Cultural thickening: intensified patterns of ties

(Couldry & Hepp, 2012)

Cultural thickening B

zh-cn
mainland China

media system 1
(nation) state 1

zh-hk
Hong Kong

media system 2
(nation) state 2

zh-tw
Taiwan

media system 3
(nation) state 3

Figure 3-2. Cultural thickening patterns that overcome or reinforce boundaries
Cultural thickening (Couldry & Hepp, 2012)

- The rise of national cultures is related to the diffusion of the so-called mass media (p.12-13).
- In this sense, media cultures are thickenings of translocal processes of meaning articulation that themselves are more (or less) locally specific (p.254).
- This means that the borders of the “cultural thickenings” to which we belong do not necessarily correspond with territorial borders, even though territories continue to be highly relevant for constructing national community (p.254).
Geo-linguistic structuring the (Chinese-language) web

- [B]oth researchers and designers can construct varied notions of web spheres as research objects (and, for designers, objects to be constructed).

- [F]actors of language and region are central in the construction of web spheres because they often serve as the baselines for sorting users into identifiable groups.
Expectation A: mainland China isolated

GFW: filtering/censorship regime

Baidu Baike

Chinese Wikipedia

zh-cn mainland China

media system 1 (nation) state 1

zh-hk Hong Kong

media system 2 (nation) state 2

zh-tw Taiwan

media system 3 (nation) state 3

Bremmer, 2006
Economist, 2013
Expectation B: a single culturally defined market

Chinese ‘culturally defined market’
(Taneja & Wu, 2013)

Baidu Baike?

Chinese Wikipedia?

zh-cn
mainland China
media system 1
(nation) state 1

zh-hk
Hong Kong
media system 2
(nation) state 2

zh-tw
Taiwan
media system 3
(nation) state 3
Expectation C: Hong Kong 1997 Handover

People's Republic of China

zh-cn
mainland China
media system 1
(nation) state 1

zh-hk
Hong Kong
media system 2
(nation) state 2

zh-tw
Taiwan
media system 3
(nation) state 3
None of the above

Baidu Baike

Chinese Wikipedia

zh-cn mainland China media system 1 (nation) state 1

zh-hk Hong Kong media system 2 (nation) state 2

zh-tw Taiwan media system 3 (nation) state 3
Main chapters and sections

- **Chapter 4**
  Editorial processes, Internet control, and Internet diffusion
  - 1) Filtering user contribution
  - 2) Gatekeeping mainland Chinese users

- **Chapter 5**
  Citation and content analysis
  - 1) Geolinguistic patterns and preferences
  - 2) Size and institutional consideration
  - 3) Defining/negotiating Chineseness

- **Chapter 6**
  Reception and use
  - 1) Search engine result pages
  - 2) Microblog posts
Editorial: filtering user contribution

Baidu Baike

Chinese Wikipedia

zh-cn mainland China
media system 1 (nation) state 1

zh-hk Hong Kong
media system 2 (nation) state 2

zh-tw Taiwan
media system 3 (nation) state 3
Figure 4-11. Chinese regions and East Asian regions: Category II average.
Baidu Baike effectively excluded users of traditional Chinese characters and censored politically sensitive content.

Chinese Wikipedia
- integrates users across regions with adequate policy and design.

Beijing used the power to *gatekeep* mainland Chinese users for user-generated websites.

Editorial processes, Internet control, and Internet diffusion
5.1 Geolinguistic patterns and preferences

5.2 Size and institutional considerations

5.3 Defining and negotiating Chineseness

Implications for showing the connections (or lack thereof) between websites and the world
Comparing distribution of links across selected regions

| Pre-normalized and normalized comparisons of Baidu Baike (BB) and Chinese Wikipedia (CW) |
|---------------------------------|---------------------------------|
| United States                  | China                           |
| Hong Kong                      | Macao                           |
| Taiwan                         | Japan                           |
| South Korea                    | Singapore                       |
| Malaysia                       | Vietnam                         |
| Australia                      | India                           |
| Brazil                         | Russia                          |
| Netherlands                    | Czech Republic                  |
| Poland                         | Iran                            |
| United Arab Emirates           | South Africa                    |

Pre-normalized (N)

<table>
<thead>
<tr>
<th>BB</th>
<th>CW</th>
</tr>
</thead>
<tbody>
<tr>
<td>cc TLD</td>
<td>cc TLD</td>
</tr>
<tr>
<td>geo-IP</td>
<td>geo-IP</td>
</tr>
</tbody>
</table>

Normalized by GDP

<table>
<thead>
<tr>
<th>BB</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>cc TLD</td>
<td></td>
</tr>
<tr>
<td>geo-IP</td>
<td></td>
</tr>
</tbody>
</table>
Comparing distribution of links across selected regions

<table>
<thead>
<tr>
<th>Normalized by GDP (N/ GDP)</th>
<th>BB</th>
<th>CW</th>
</tr>
</thead>
<tbody>
<tr>
<td>cc TLD</td>
<td></td>
<td></td>
</tr>
<tr>
<td>geo-IP</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cc TLD</td>
<td></td>
<td></td>
</tr>
<tr>
<td>geo-IP</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Normalized by Internet population (N/ pop internet)</th>
<th>BB</th>
<th>CW</th>
</tr>
</thead>
<tbody>
<tr>
<td>cc TLD</td>
<td></td>
<td></td>
</tr>
<tr>
<td>geo-IP</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cc TLD</td>
<td></td>
<td></td>
</tr>
<tr>
<td>geo-IP</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Definitions

Search engine result pages (SERPs)
- Baidu and Baidu Baike: mainland China.
- Chinese Wikipedia: visible across regions

Microblog posts
- Both Verbification: c.f. Google(v) sth.
- Baidu Baike: censored and of low quality
- Chinese Wikipedia: free and reliable

Implications for information engagement

Findings

Reception and use

Chapter 6

6.1 Search engine result pages (SERPs)

6.2 Microblog posts
### Percentage of visibility scores: encyclopedia sites among the top-10

<table>
<thead>
<tr>
<th>Categories</th>
<th>Baidu CN</th>
<th>Yahoo CN</th>
<th>Google CN</th>
<th>Google SG</th>
<th>Yahoo SG</th>
<th>Google HK</th>
<th>Google TW</th>
<th>Yahoo HK</th>
<th>Yahoo TW</th>
</tr>
</thead>
<tbody>
<tr>
<td>The Cambridge Encyclopedia of China</td>
<td><img src="image1.png" alt="Graph" /></td>
<td><img src="image2.png" alt="Graph" /></td>
<td><img src="image3.png" alt="Graph" /></td>
<td><img src="image4.png" alt="Graph" /></td>
<td><img src="image5.png" alt="Graph" /></td>
<td><img src="image6.png" alt="Graph" /></td>
<td><img src="image7.png" alt="Graph" /></td>
<td><img src="image8.png" alt="Graph" /></td>
<td><img src="image9.png" alt="Graph" /></td>
</tr>
<tr>
<td>Top 10 Search Terms (Google and Baidu)</td>
<td><img src="image1.png" alt="Graph" /></td>
<td><img src="image2.png" alt="Graph" /></td>
<td><img src="image3.png" alt="Graph" /></td>
<td><img src="image4.png" alt="Graph" /></td>
<td><img src="image5.png" alt="Graph" /></td>
<td><img src="image6.png" alt="Graph" /></td>
<td><img src="image7.png" alt="Graph" /></td>
<td><img src="image8.png" alt="Graph" /></td>
<td><img src="image9.png" alt="Graph" /></td>
</tr>
<tr>
<td>Best Film/Popular Music (China, Hong Kong, Taiwan)</td>
<td><img src="image1.png" alt="Graph" /></td>
<td><img src="image2.png" alt="Graph" /></td>
<td><img src="image3.png" alt="Graph" /></td>
<td><img src="image4.png" alt="Graph" /></td>
<td><img src="image5.png" alt="Graph" /></td>
<td><img src="image6.png" alt="Graph" /></td>
<td><img src="image7.png" alt="Graph" /></td>
<td><img src="image8.png" alt="Graph" /></td>
<td><img src="image9.png" alt="Graph" /></td>
</tr>
<tr>
<td>Modern Concepts</td>
<td><img src="image1.png" alt="Graph" /></td>
<td><img src="image2.png" alt="Graph" /></td>
<td><img src="image3.png" alt="Graph" /></td>
<td><img src="image4.png" alt="Graph" /></td>
<td><img src="image5.png" alt="Graph" /></td>
<td><img src="image6.png" alt="Graph" /></td>
<td><img src="image7.png" alt="Graph" /></td>
<td><img src="image8.png" alt="Graph" /></td>
<td><img src="image9.png" alt="Graph" /></td>
</tr>
</tbody>
</table>
Percentage of visibility scores: encyclopedia

Modern Concepts
(shared with modern Japanese)

Notable People

Potentially sensitive terms

Fortune500

- baidu.com
- wikipedia.org
- hudong.com
- mbalib.com
Patterns Found

Table 7-1

<table>
<thead>
<tr>
<th>Definitions</th>
<th>Findings</th>
<th>Conclusions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baidu Baike</td>
<td>Chinese Wikipedia</td>
<td></td>
</tr>
<tr>
<td><strong>is (relatively more) open to</strong></td>
<td>copyright-dubious, self-promoting, or advertisement-based content</td>
<td>reliable information sources across Chinese-speaking regions</td>
</tr>
<tr>
<td><strong>is (relatively more) defensive against</strong></td>
<td>topics politically sensitive to Beijing and information sources outside mainland China</td>
<td>copyright-dubious, self-promoting, or advertisement-based content</td>
</tr>
<tr>
<td><strong>and thus produces cultural thickening</strong></td>
<td>commercially and politically correct mainland China-focused Chinese information order</td>
<td>culturally and politically diverse Chinese information order across various Chinese societies</td>
</tr>
</tbody>
</table>
Views on the impact of GFW (See Table 7-2)

- **Isolation:**
  strong GFW isolates mainland China from the world

- **Indigenization:**
  GFW does not prevent the formation of a single Chinese cluster

- **‘Protection’** of mainland China from the rest of the (Chinese-speaking) world
  - Hong Kong and Taiwan as in-between nodes
  - Beijing creates comparative advantage in providing access to mainland Chinese users commercially, culturally and politically
  - GFW lowers the intensity of exchange between the majority of mainland Chinese users and the rest
It is a struggle to “centre” the Chinese-language web sphere by promoting certain directions, intensities and normative values via cultural thickening processes among new internet users.
Research Implications

More research on geolinguistic factors and/or processability is needed:

The Web is ‘increasingly grounded with geographic and linguistic specificity by platform and space’ (R. Rogers, 2013, p.58)

Practical Implications

Website designers need to understand the geolinguistic profile of their users

Users need to be aware of the limits imposed

Policy Implications

Improve the cultural politics of web spaces
Beyond PhD

- Special Issue on “Social Data Science and the Chinese Web” for *Policy and Internet*

- Visualizing interlanguage links of all Wikipedia
  - Leonardo: Art, Science and Technology: [http://goo.gl/N9zIAd](http://goo.gl/N9zIAd)
  - More: [http://goo.gl/6ZI7tz](http://goo.gl/6ZI7tz)

- Visualizing geolinguistic traffic of Wikipedia

- Intensity of cyber crimes and attacks: How geographic normalization facilitates cross-country comparison
  - IEEE Security & Privacy

- Internet diffusion in Chinese-speaking regions, Asia, and the world
Figure 1. The trendlines

growth: 2002-2011

decline?
a growing number of publications

- Figure 1
- Growth since 2002
- a two- to three-year window
  - where major academic databases index fresh research.
- The Hong Kong results are identical to the Taiwanese ones.
Major (archive) platforms

Table 3. Most frequently appearing source platforms

<table>
<thead>
<tr>
<th>Source Platform</th>
<th>CNKI Scholar</th>
<th>Google Scholar China</th>
<th>Google Scholar Taiwan</th>
<th>Grand Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>CNKI platform</td>
<td>464</td>
<td>219</td>
<td>0</td>
<td>683</td>
</tr>
<tr>
<td>Education Taiwan</td>
<td>0</td>
<td>2</td>
<td>605</td>
<td>607</td>
</tr>
<tr>
<td>Miscellaneous</td>
<td>0</td>
<td>346</td>
<td>248</td>
<td>594</td>
</tr>
<tr>
<td>CQVIP platform</td>
<td>0</td>
<td>378</td>
<td>0</td>
<td>378</td>
</tr>
<tr>
<td>Airitilibrary platform</td>
<td>0</td>
<td>0</td>
<td>147</td>
<td>147</td>
</tr>
<tr>
<td>Wanfang Data platform</td>
<td>0</td>
<td>55</td>
<td>0</td>
<td>55</td>
</tr>
<tr>
<td>Grand Total</td>
<td>464</td>
<td>1000</td>
<td>1000</td>
<td>2464</td>
</tr>
</tbody>
</table>
Cross tab analysis: Major publication institutions

Table 2. Most frequently appearing source institutions

<table>
<thead>
<tr>
<th>Institutions (in Chinese)</th>
<th>Institutions</th>
<th>CNKI Scholar</th>
<th>Google Scholar China</th>
<th>Google Scholar Taiwan</th>
<th>Grand Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>情报学报</td>
<td>Journal of the China Society for Scientific and Technical Information</td>
<td>0</td>
<td>65</td>
<td>0</td>
<td>65</td>
</tr>
<tr>
<td>国家图书馆学刊</td>
<td>Journal of the National Library of China</td>
<td>1</td>
<td>41</td>
<td>0</td>
<td>42</td>
</tr>
<tr>
<td>臺灣大學企業管理碩士專班學位論文</td>
<td>MBA dissertations, National Taiwan University</td>
<td>0</td>
<td>0</td>
<td>29</td>
<td>29</td>
</tr>
<tr>
<td>臺灣師範大學工業科技教育學系學位論文</td>
<td>Department of Industrial Technology Education Dissertations, Taiwan Normal University</td>
<td>0</td>
<td>0</td>
<td>28</td>
<td>28</td>
</tr>
<tr>
<td>圖文傳播藝術學報</td>
<td>NTUA Department of Graphic Communication Arts</td>
<td>0</td>
<td>0</td>
<td>23</td>
<td>23</td>
</tr>
<tr>
<td>现代图书馆</td>
<td>Modern Technology of Library and Information</td>
<td>5</td>
<td>12</td>
<td>0</td>
<td>17</td>
</tr>
<tr>
<td>互联网周刊</td>
<td>China Internet Weekly</td>
<td>6</td>
<td>10</td>
<td>0</td>
<td>16</td>
</tr>
<tr>
<td>图书情报工作</td>
<td>Library and Information Service</td>
<td>5</td>
<td>9</td>
<td>0</td>
<td>14</td>
</tr>
<tr>
<td>青年记者</td>
<td>&quot;Youth Reporter&quot; magazine</td>
<td>3</td>
<td>11</td>
<td>0</td>
<td>14</td>
</tr>
<tr>
<td>中文信息学报</td>
<td>Journal of Chinese Information Processing</td>
<td>6</td>
<td>7</td>
<td>0</td>
<td>13</td>
</tr>
<tr>
<td>互联网天地</td>
<td>China Internet Magazine</td>
<td>7</td>
<td>6</td>
<td>0</td>
<td>13</td>
</tr>
<tr>
<td>上海交通大学</td>
<td>Shanghai Jiao Tong University</td>
<td>13</td>
<td>0</td>
<td>0</td>
<td>13</td>
</tr>
</tbody>
</table>
Cross tab analysis: Major publication institutions

(continued)

- Shanghai Jiao Tong University
  Information and Documentation Services 3 9 0 12
  Performance Arts Review 0 0 12 12
  Library Theory and Practice 5 6 0 11
  Information Studies: Theory & Application 5 6 0 11
  Harbin Institute of Technology
    Journal of Intelligence 7 3 0 10
    Modern Information 2 8 0 10
  上海交通大学
  情报资料工作
  表演藝術雜誌
  图书馆理论与实践
  情报理论与实践
  哈尔滨工业大学
  情报杂志
  现代情报

Between CNKI Scholar and Google Scholar China, substantial overlapping results exists for journals and magazines. Google Scholar China does not seem to include as many degree dissertations as CNKI Scholar.
何舟: 新闻与传播研究发表机制

- 台湾的发表机制与美国的近似。由于有一定数量的研究人才和一定规模的学术社区，台湾形成了一个自成体系的、以中文为主的发表机制。岛内的学者绝大部分除偶尔在国际期刊上发表外，一般都在本土的学术期刊和出版机构发表自己的研究成果。香港的学者由于社区很小，没有自己的刊物，再加上各大学一般只承认以英文在国际期刊上发表的文章，所以绝大部分的研究成果都在美国或欧洲的学术刊物上发表。中国大陆的发表机制在相当长的时间里并没有采取专家匿名评审的方式，只是在近年来逐渐开始实行这种方式。由于大陆有广大的市场和学术社区，大陆的研究成果绝大部分以中文在本土发表，绝少在国际刊物上以英文或其他文字发表。
1. Methodology and Research Strategy

2. Digital methods:
   … reintroducing quantitative methods and formal system analysis into the field of political/social sciences

3. Big state, Big data industry and research interventions:
   … researchers must reflect on the methods, data and, arguably most important of all, possible interventions in the daily routine of online interactions, social association, and offline people movements.
Big Data as Data centralism

- Such work will encourage the reflections on the existing political principle of “democratic centralism” (minzhuzhongzhi 民主集中制) in relation to governing techniques and technologies.
- The relationship between “democratic centralism” and “data centralism”
Big Data

- Herrenhausen Conference on Big Data, on March 25-27 2015 in Hannover, Germany.

- Q: in what ways have the big data industry of online expressions ... make their respective information and communication spaces towards or away from the desired values of democratic public sphere?
My arguments

- (1) Big data analysis is often large-scale data analysis searching for meaningful reductions.
- (2) Big data industry of online expressions often scale up and then reduce online expressions to construct information and communication spaces for both social and personal meanings.
- (3) According to Lewis Mumford, the tensions “between small-scale association and large-scale organization, …, between remote control and diffused local interventions” have been major challenges for democratic association to scale up from small communities and groups.
  - (3a) Democratic technics are thus “small scale method of production, … even when employing machines, remaining under the active direction of the craftsman. They are man-centered. relatively weak. but resourceful
My arguments

- (3b) Authoritarian technics are often “system-centered, immensely powerful, but inherently unstable”
- (4) Social scientists and theorists must render big data technics of online expressions more democratic than authoritarian.
- (5) They also must develop distributive methods that can intervene in the ways in which the big data industry manage online expressions
WHY IT MATTERS: INTERNATIONALIZATION OF INTERNET RESEARCH
Major five world languages: potentials for scaling up online expressions processability

- **Mandarin:**
  - Simplified Chinese (zh-Hans), official in China (zh-CN) and Singapore (zh-SG); Traditional Chinese (zh-Hant), official in Taiwan (zh-TW), Hong Kong (zh-HK) and Macau (zh-MO)

- **Spanish**
  - Official in Mexico (es-MX), Spain (es-ES), Argentina (es-AR), Colombia (es-CO), etc.

- **English:**
  - Official in US (en-US), India (en-IN), Nigeria (en-NG), Pakistan (en-PK), Philippines (en-PH), UK (en-UK), etc.

- **Hindi:** Official in India (hi-IN)

- **Bengali:** Official in Bangladesh (bn-BD) and India (bn-IN)

- **Arabic:**
  - Official in Egypt (ar-EG), Algeria (ar-DZ), Saudi Arabia (ar-SA), Iraq (ar-IQ), Sudan (ar-SD), Yemen (ar-YE)

- **Russian:**
  - Official in Russia (ru-RU), Ukraine (ru-UA), Kazakhstan (ru-KZ), Kyrgyzstan, Belarus (ru-BY)
Major five world languages: consider the polities and internet companies involved

- Mandarin: the most intriguing case (BAT in zh-CN)
  - All digitally advanced with different free/unfree and/or democratic/authoritarian polities
- Spanish
  - Official in Mexico (es-MX), Spain (es-ES), Argentina (es-AR), Colombia (es-CO), etc.
- English:
  - Official in US (en-US), India (en-IN), Nigeria (en-NG), Pakistan (en-PK), Philippines (en-PH), UK (en-UK), etc.
- Hindi: Official in India (hi-IN)
- Bengali: Official in Bangladesh (bn-BD) and India (bn-IN)
- Arabic:
  - Official in Egypt (ar-EG), Algeria (ar-DZ), Saudi Arabia (ar-SA), Iraq (ar-IQ), Sudan (ar-SD), Yemen (ar-YE)
- Russian: (Yandex in ru-RU)
  - Official in Russia (ru-RU), Ukraine (ru-UA), Kazakhstan (ru-KZ), Kyrgyzstan (ru-KG), Belarus (ru-BY)
Research Agenda:

We need more research on web spheres in relation to the concepts of a (Chinese) cultural sphere, national sphere, and transnational public sphere.
Thank you