Brief analysis of questionnaire survey on recognition and use of Japanese social science data

2020-3-19

Yasuyuki Minamiyama, Chief Data Librarian
Research Center for Open Science and Data Platform (RCOS),
National Institute of Informatics (NII)
Questionnaire survey on recognition and use of Japanese social science data

The objective of this survey is to assess the awareness and usage of Japanese social science data in North America. Through this survey, we aim to get some insights into the needs for Japanese social science data in North America and to understand the issues in secondary use from outside Japan. The results of the survey will be presented at the “Council on East Asian Libraries Annual Meeting: Boston, Massachusetts, CJM Program” held on March 19, 2020 and may be published in an appropriate journal. Thank you for your cooperation!

<Information>

Period: Feb. 12, 2020 to Mar. 4, 2020 (EST)
Outline: 19 questions／about 20 minutes
Contact: Yasuyuki Minamiyama, Chief Data Librarian, Research Center for Open Science and Data Platform, National Institute of Informatics (minamiyama@nii.ac.jp)
1. General information

Respondent's institution size

All universities less than 40,000
Not included any research institute
Types of research data services

Other:
* Assist faculty and students to plan "Research data cycle"

* We don't provide all those services at departmental Library, but we do as the whole university library

* Providing information for the entire library system
### Years/staff members

#### Years involved

```
<table>
<thead>
<tr>
<th>Years involved</th>
<th>1 to 3 year</th>
<th>3 to 5 year</th>
<th>5 and more</th>
<th>I don't know</th>
</tr>
</thead>
<tbody>
<tr>
<td>7,000 – 9,999</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>10,000 – 19,999</td>
<td></td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>20,000 – 29,999</td>
<td></td>
<td></td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>30,000 – 39,999</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>40,000 – 99,999</td>
<td></td>
<td>2</td>
<td></td>
<td>1</td>
</tr>
</tbody>
</table>
```

#### Number of Staff

```
<table>
<thead>
<tr>
<th>Years involved</th>
<th>1 to 3</th>
<th>4 to 6</th>
<th>7 to 9</th>
<th>10 and more</th>
<th>I don't know</th>
</tr>
</thead>
<tbody>
<tr>
<td>7,000 – 9,999</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>10,000 – 19,999</td>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>20,000 – 29,999</td>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>30,000 – 39,999</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>40,000 – 99,999</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>
```

Libraries for research data services are no longer new!
Types of research data (in social sciences)

- Geospatial data
- Text data
- Survey data
- Aggregate data
- Other individual data (in social sciences)
- I don't know
- SNS data (e.g. twitter)
- Experimental data
- None

N=11

* Geospatial data
* Text data
* Survey data
* Aggregate data

may be often covered
2. About dealing with Japanese social science data
Data sources (by website type)

- disciplinary data repositories/data archives
- academic institutional repositories/data archives

...they not be used much
Data sources (by country)

United States and Japan only
Any concerns about using data retrieved and obtained

- Difficult to use data format
- No concerns
- Information on obtaining procedures is available only in Japanese
- Data scattered
- Other
- How to use
- Unclear procedures for obtaining data
- Unable to search for data by specifying detailed conditions
- Data quality is poor
- Do not understand the meaning of the data (too complicated)
- Incomplete or insufficient documentation
- Data statistics and analysis methods
- Data visualization / graphing / presentation material creation
- Do not understand the usage conditions of data
- Data download takes a long time / slow search speed
- Data does not conform to standards

Other:
* “It is very discouraging to find data related to Japan. So many users would quit searching or do not bother from the very beginning.”

Number
3. About publishing Japanese social science data
Preparation status of institutional data repository

- Yes: 81.8%
- No: 9.1%
- I do not know: 9.1%

N=11
Data repository recommended experience

Only 2 institutions have experience (Number of requests: 4 in total).
Expected properties of the data repositories/data archives

Sustainability (continuation of related projects / organizations)
Stability (resistant to network failures, power outages, disasters, etc.)
International (use across national boundaries is possible)
Cost reduction (whether it will reduce your work and financial burden)
Interdisciplinary (can be used across discipline)
I do not know
Guaranteed data quality reliability
Other
Familiarity / authority (famous internationally or in the discipline)
Commonality (non-profit, etc.)
Expertise (specialization for use within the discipline)

Other:
* Centralized manage infrastructure

N=11
Most important features of the data repositories/data archives

- Search function (easy to find data)
- Quality control function (guaranteed data quality)
- Ability as a library (covers a lot of data and information)
- Download function (easy data acquisition)
- Archive capacity (with sufficient data capacity)
- Sharing function (fine access restrictions and terms of use can be set)
- Publishing function (easy for any user to use)
- Other
- Analysis function (models are substantial)
- Upload function (easy to provide data)
Recognitions of Japanese Data Repositories / Aggregation Services / Government Sites

N=10

- KAKEN Grants
- J-STAGE
- CiNii Articles/Books/Dissertations
- Institutional Repositories Database (IRDB)

N=10

- Statistics Bureau of Japan
- data.go.jp
- e-Stat
- e-Gov
Additional comments

• Japanese social data are scattered. It is extremely important to have ONE search (metadata of datasets) to reach each institutional data set. I know that Japanese government has already taken this project, and I really looking forward to the completion of this project.

  I totally agree your opinion. Many thanks for your expectation!

• I would love to be able to access the questions in this survey so I could study some of the services and databases listed here.

  Please see the link below:
  https://filesender.nii.ac.jp/?vid=3e295dd7-be08-9688-2307-00007ef4fc5f

• Experience with Japanese social science data is mostly limited to statistical information

  As far as I have heard, Japanese researchers are in a similar situation. The situation may change when secondary use of qualitative data begins.
Thank you!

Special thanks (honorifics omitted):

• Fabiano Rocha
• CJM members (Hamish Todd, Liu Wen-ling, Sachiko Iwabuchi, Setsuko Noguchi, Azusa Tanaka, Toshie Marra, and more?)
• And all of the respondents!