Data Literacy Perceptions and Research Data Management Practices by Researchers in Japan

Ui IKEUCHI\textsuperscript{1)*}, Takashi HARADA\textsuperscript{2)}, Sho SATO\textsuperscript{2)}, Yukinori OKABE\textsuperscript{2)}, and Hiroshi ITSUMURA\textsuperscript{1)}

\textsuperscript{1)University of Tsukuba, Japan; \textsuperscript{2)Doshisha University, Japan

*oui@slis.tsukuba.ac.jp / ikeuchi.ui@gmail.com
http://orcid.org/0000-0002-5680-1881
University of Tsukuba

National Research University

http://www.tsukuba.ac.jp/en/
Doshisha University

Private Research University

26 Libraries
29,459 students
Contents

1. Introduction – Researchers in Japan
2. Methods
3. Respondents Overview
4. Results & Discussion
5. Conclusion

❖ Acknowledgements
1. Researchers in Japan
R&D expenditure [1981-2014]

Trillion, JPY

YEAR

2001

NISTEP. Science and Technology Indicators 2016
Research Time

Ideal: 46.9%
- Research: 12.1%
- Education: 29.3%
- Social Service: 11.7%
- others: 46.9%

Reality: 35.0%
- Research: 17.5%
- Education: 28.4%
- Social Service: 19.1%
- others: 35.0%

GAP

 Declining article productivity

<table>
<thead>
<tr>
<th>Year</th>
<th>Policy Title</th>
<th>Author/Proponent</th>
</tr>
</thead>
<tbody>
<tr>
<td>2015.3</td>
<td>Increasing Scientific Research Transparency</td>
<td>(Science Council of Japan)</td>
</tr>
<tr>
<td>2015.3</td>
<td>Promoting Open Science in Japan: Opening up a new era for the advancement of science</td>
<td>(Expert Panel on Open Science, based on Global Perspectives Cabinet Office)</td>
</tr>
<tr>
<td>2016.1</td>
<td>5th Science and Technology Basic Plan: 2016-2020</td>
<td>(Cabinet Office, Government of Japan)</td>
</tr>
<tr>
<td>2016.2</td>
<td>Data Management Policy for Strategic Basic Research Programs</td>
<td>(Japan Science and Technology Agency; JST)</td>
</tr>
<tr>
<td>2016.7</td>
<td>Promoting Open Access and Open Research Data</td>
<td>(Ministry of Education, Culture, Sports, Science and Technology)</td>
</tr>
<tr>
<td>2017.4</td>
<td>JST Policy on Open Access to Research Publications and Research Data Management</td>
<td>(JST Open Science Policy)</td>
</tr>
</tbody>
</table>
Science Council of Japan (2015)

*Increasing Scientific Research Transparency*

– Measures against scientific misconduct

Preserve research data for 10 years

PI is required to develop a data management plan (DMP)

DMP should be submitted before the research project begins
Increasing the total number of papers

Increasing the top 10% most-often cited papers

Promoting open science

http://www8.cao.go.jp/cstp/english/basic/5thbasicplan.pdf
Researchers in Japan

Preserve data

DMP

Open Science
Open data
Purpose

To support researchers in Japan

Develop RDM (Research Data Management) service

– Appropriate data literacy training
– Institutional [Data] Repository
Research Questions

1. Does researchers have open data experience?

2. Does researchers have concerns about data sharing? What is the most concern?

3. Does researchers would like to have a formal training? What is the most interesting thing?
2. Methods
JP Data Literacy Survey

You are invited to participate in a survey which aims to collect data about the data literacy of academics and research students in higher education institutions. From your responses we will be able to fully understand the current levels of awareness and gaps in knowledge which will help us develop appropriate data literacy training for the higher education community.

Please answer all the questions, and note that this survey is anonymous. It will take approximately 20 minutes to complete the entire survey.

By completing this survey you are consenting to the use of your data for research and dissemination purposes. If you have any questions or comments as you are going through the survey, please contact Ui Ikeuchi [oui@slis.tsukuba.ac.jp]*.

Thank you very much for your cooperation!

Survey Period

- 24 February - 17 March 2017
Questionnaire (Japanese)

Added explanations about data policies in Japan

“Respondents without experience of RDM or Open Research Data are also welcome.”
19 [] 以下の質問について、あてはまるものをお選び下さい。

* 各選択肢について、適切な回答を選択してください。

所属大学にはデータ管理計画（Data Management Plan: DMP）がありますか？（研究に用いるデータの種類・管理・公開・保存方法などを記載した計画書のことで、大学や助成機関に提出する。詳しくは「研究データ管理 (http://id.nii.ac.jp/1280/00000195/)」参照）
Email Invitation

Sending from

– University of Tsukuba Library
– Department of Research Promotion, Univ. of Tsukuba
– Organization for Research Initiatives and Development, Doshisha University
– The authors
Survey Outline

- Respondents
  - Academic Staff in Japan
  - Research Student (Ph.D Student) in Japan

- Period
  - 24 February - 24 March 2017

*The University of Tsukuba approved the Research Ethics Application for this study on 4 January 2017 (No. 28-121).
3. Respondents Overview
Valid responses & Current role

Academic Staff: 47.3%
Research Student: 52.7%
Age (n=586)

<table>
<thead>
<tr>
<th>Age Group</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>18-25</td>
<td>3.1%</td>
</tr>
<tr>
<td>26-35</td>
<td>41.0%</td>
</tr>
<tr>
<td>36-45</td>
<td>23.7%</td>
</tr>
<tr>
<td>46-55</td>
<td>19.8%</td>
</tr>
<tr>
<td>56-65</td>
<td>11.1%</td>
</tr>
<tr>
<td>65+</td>
<td>0.9%</td>
</tr>
</tbody>
</table>

Don’t want to disclose | 0.5%
Discipline (n=586)

- Social Sciences: 26.6%
- Natural Sciences: 23.5%
- Medical & Health Sciences: 20.3%
- Humanities: 15.9%
- Engineering & Technology: 13.1%
- Others: 0.5%
4. Results & Discussion
Q16 Do you collaborate with other researchers and share data?

65.7% of n=586 respondents indicated they do.
Q16 Do you collaborate with other researchers and share data?

- Engineering & Technology: 81.8%
- Natural Sciences: 78.3%
- Medical & Health Sciences: 73.9%
- Social Sciences: 55.8%
- Humanities: 40.9%
- Others: 33.3%

n=586, $\chi^2 = 55.817$, $p = 0.000$
Q17 Which of the following applies to your research data?

My data is available under a variety of conditions

73.5% (n=586)
Q17 Which of the following applies to your research data?

Significant difference depending on the discipline

<table>
<thead>
<tr>
<th>Discipline</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Engineering &amp; Technology</td>
<td>94.6%</td>
</tr>
<tr>
<td>Natural Sciences</td>
<td>89.8%</td>
</tr>
<tr>
<td>Medical &amp; Health Sciences</td>
<td>85.5%</td>
</tr>
<tr>
<td>Social Sciences</td>
<td>64.5%</td>
</tr>
<tr>
<td>Humanities</td>
<td>51.7%</td>
</tr>
<tr>
<td>Others</td>
<td>66.7%</td>
</tr>
</tbody>
</table>

Average: 76.6%

n=563, $\chi^2 = 74.384$, $p = 0.000$
Q17 Which of the following applies to your research data?

- My data is available under a variety of conditions: 73.5%
- My data is openly available to everyone: 6.7%
- My data is not available to anyone else: 32%
Q18 Do you have any concerns for sharing data with others?

74.6%

n=586
<table>
<thead>
<tr>
<th>Concern</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Misinterpretation of data</td>
<td>39.1%</td>
</tr>
<tr>
<td>Fear of losing the scientific edge</td>
<td>38.6%</td>
</tr>
<tr>
<td>Misuse of data</td>
<td>37.7%</td>
</tr>
<tr>
<td>Legal and ethical issues</td>
<td>36.0%</td>
</tr>
<tr>
<td>Lack of appropriate policies and rights protection</td>
<td>24.1%</td>
</tr>
<tr>
<td>Lack of resources</td>
<td>14.3%</td>
</tr>
</tbody>
</table>
Q18 Do you have any concerns for sharing data with others?

<table>
<thead>
<tr>
<th>Concern</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Misinterpretation of data</td>
<td>39.1%</td>
</tr>
<tr>
<td>Fear of losing the scientific edge</td>
<td>38.6%</td>
</tr>
<tr>
<td>Misuse of data</td>
<td>37.7%</td>
</tr>
<tr>
<td>Legal and ethical issues</td>
<td>36.0%</td>
</tr>
<tr>
<td>Lack of appropriate policies and rights protection</td>
<td>24.1%</td>
</tr>
<tr>
<td>Lack of resources</td>
<td>14.3%</td>
</tr>
</tbody>
</table>
Q18 Do you have any concerns for sharing data with others?

- Misinterpretation of data: 39.1%
- Fear of losing the scientific edge: 38.6%
- Misuse of data: 37.7%

✓ Appropriate (interoperable) metadata
✓ Article links using permanent ID (i.e. DOI)
✓ Author(s) links using popular ID (i.e. ORCID)
Q18 Do you have any concerns for sharing data with others?

- Misinterpretation of data: 39.1%
- Fear of losing the scientific edge: 38.6%
- Misuse of data: 37.7%
- Legal and ethical issues: 36.0%
- Lack of appropriate policies and rights protection: 24.1%
- Lack of resources: 14.3%

n=586
Q18 Do you have any concerns for sharing data with others?

- Misinterpretation of data: 39.1%
- Fear of losing the scientific edge: 38.6%
- Misuse of data: 37.7%

✓ Research ethics education
✓ Data citation guidance
✓ Timestamp system
Q25 Would you like to have a formal training on the following?

73.9%  

n=586
Q18 concern(s) & Q25 training

<table>
<thead>
<tr>
<th>Q25 Would like to have a formal Training</th>
<th>Q18 Have concern(s) for sharing data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>Yes: 59.1%</td>
</tr>
<tr>
<td></td>
<td>No: 17.6%</td>
</tr>
<tr>
<td>No</td>
<td>Yes: 15.2%</td>
</tr>
<tr>
<td></td>
<td>No: 8.1%</td>
</tr>
</tbody>
</table>

n=567, $\chi^2 = 7.450, p < 0.01$
Q25 Would you like to have a formal training on the following?

- Data Management Plan: 61.4%
- Consistent file naming: 47.8%
- Metadata: 44.9%
- Version control: 42.8%
- Data citation styles: 41.0%

n=586
5. Conclusion
RQ 1 & Results

Does researchers have open data experience?
– Only 6.7%

New Question
– How can RDM service contribute to Open Research Data?
RQ 2 & Results

- Does researchers have concerns about data sharing?  
  - 74.6%

- What is the most concern?  
  - Misinterpretation (also Losing the science edge, Misuse, and...)

44
RQ 3 & Results

- Does researchers would like to have a formal training?
  - 73.6%

- What is the most interesting thing?
  - DMP (Data Management Plan)
Suggestions for RDM service

- Appropriate data literacy training
  - Provide DMP, research ethics, and data citation guidance
  - Consider the differences among disciplines

- Institutional [Data] Repository
  - Metadata, DOI, and researchers’ ID
  - Timestamp

Support researchers’ RDM activities and reduce concerns about data sharing
Future work

Staff development for RDM service

- i.e. Research Data Alliance
  
  ✓ Education and Training on handling of research data IG (Interest Group)
  
  ✓ Libraries for Research Data IG

http://rd-alliance.org
Target: Research supporters

- Librarian, IT specialist, URA (University Research Assistant), ...

Nov. 2017-

https://lms.gacco.org/courses/course-v1:gacco+ga088+2017_11/about
Acknowledgements

University of Tsukuba
– Center for Artificial Intelligence Research
– Department of Research Promotion
– University of Tsukuba Library
– Overseas Academic Conference Participation Support Program

Doshisha University
– Organization for Research Initiatives and Development