

## **Data supporting “Investigation and Development of the workflow to clarify conditions of use for research data publishing in Japan” Version 1.0**

Last updated 2020-11-26 by Yasuyuki Minamiyama

### 1. Author Information

Name: Ui Ikeuchi (Data collector)

Institution: Department of English Language & Literature, Faculty of Language & Literature, Bunkyo University, Saitama, Japan

Email: [ikeuchi@bunkyo.ac.jp](mailto:ikeuchi@bunkyo.ac.jp)

ORCID: <https://orcid.org/0000-0002-5680-1881>

Name: Yasuyuki Minamiyama (Data curator)

Institution: Research Center for Open Science and Data Platforms, National Institute of Informatics, Tokyo, Japan

Email: [minamiyama@nii.ac.jp](mailto:minamiyama@nii.ac.jp)

ORCID: <http://orcid.org/0000-0002-7280-3342>

2. Date of data collection: 2018-02-13 to 2018-03-20

3. Geographic location of data collection: Japan

### 4. Data Description

This survey aims to clarify the actual situation and perceptions of using the conditions of use through a questionnaire survey for data holders. At the same time, we also aim to obtain a more specific knowledge of giving incentives to data holders, which have already been pointed out. Ten questions are provided, and no mandatory items are included. Some of the questions are expected to be difficult for some respondents to answer.

The survey period was set from February 13, 2018 to March 20, 2018. A web questionnaire was distributed via some mailing lists and websites using Questant's questionnaire system. The survey not only mainly targeted researchers, but also data manager, and librarians. The final number of responses is 413, of which 409 are valid responses. The aggregated results of the 409 valid responses are presented in the paper.

## 5. List of Variables

NA (無回答) signifies that the information was not available. "blank" written out in a cell signifies that the response was left blank by the survey taker.

List of Variables

Variable Name	Description	Type	Value Labels	Value Labels (jp)
SAMPLEID	respondents ID	Numeric		
ANSWERDATE	Date and time of data collection	Date, yyyy/mm/dd tt:tt		
CELL	Group number of respondents	Numeric		
CELLNAME	Name of group	String		
Q1	Property of respondents	Numeric	1="I am not currently engaged in any research activities"	1="現在、研究活動は行っていない"
			2="Engineering"	2="工学"
			3="Astronomy"	3="天文学"
			4="Chemistry"	4="化学"

			5="Physics"	5="物理学"
			6="Earth Sciences"	6="地球科学"
			7="Mathematics"	7="数学"
			8="Computer Science"	8="コンピュータサイエンス"
			9="Agricultural Science"	9="農学"
			10="Biological Science"	10="生物科学"
			11="Medicine"	11="医学"
			12="Psychology"	12="心理学"
			13="Social Sciences"	13="社会科学"
			14="Humanities"	14="人文学"
			15="Other"	15="その他"
Q1_15FA	Property of respondents (15:Other)	String		
Q2S1_1 to Q2S9_3	Experience in obtaining published data and publishing data by yourself	Numeric	0="No"	0="いいえ"
			1="Yes"	1="はい"
			NA="No Answer"	NA="無回答"
Q3S1 to Q3S3	Awareness of existing licenses	Numeric	1="Yes"	1="はい"
			2="No"	2="いいえ"
			NA="No Answer"	NA="無回答"
Q4S1 to Q4S3	Usage of existing licenses	Numeric	1="Yes"	1="はい"

			2="No"	2="いいえ"
			*=skip logic	*=スキップ
			NA="No Answer"	NA="無回答"
Q5S1 to Q5S11	Desired condition of use when respondents publish their research data	Numeric	1="Yes"	1="はい"
			2="It depends on cases"	2="場合による"
			3="No"	3="いいえ"
			4="I don't know"	4="わからない"
			NA="No Answer"	NA="無回答"
Q6	License compliance and willingness to publish data	Numeric	1="Agree"	1="公開してもよいと思う"
			2="Somewhat agree"	2="やや公開してもよいと思う"
			3="Somewhat disagree"	3="あまり公開したくないと思う"
			4="Disagree"	4="公開したくないと思う"
			5="I don't know"	5="わからない"
Q7_1 to Q7_6	An appropriate method of displaying the use of published data	Numeric	0="No"	0="いいえ"
			1="Yes"	1="はい"
			NA="No Answer"	NA="無回答"
Q7_5FA	An appropriate method of displaying the use of	String		

	published data (5:Other)			
Q81FA	Requests and concerns about data reuse	String		
Q9_1 to Q9_8	Desired approach to data use and publishing	Numeric	0="No"	0="いいえ"
			1="Yes"	1="はい"
			NA="No Answer"	NA="無回答"
Q9_7FA	Desired approach to data use and publishing (7:Other)	String		

## 6. List of Questions

Q1. Which of the following terms best describes your research field?

Q2. Have you ever obtained or published any data, including the cases in which user registration and fees are required? (Choices: Obtain="\_1", Publish="\_2", None="\_3")

- Personal/research lab websites or blogs (See Q2S1\_1-3)
- Institutional repositories/data archives (See Q2S2\_1-3)
- Data repositories/archives in specific field (See Q2S3\_1-3)
- Government repositories/data archives (See Q2S4)
- Repositories/data archives by Commercial company (See Q2S5)
- Other data publishing services (e.g. figshare, zenodo) (See Q2S6\_1-3)
- Code sharing services (e.g. GitHub) (See Q2S7\_1-3)
- Supplementary materials (in research paper) (See Q2S8\_1-3)
- Academic SNS services (e.g. Mendeley, ResearchGate) (See Q2S9\_1-3)

Q3. Are you familiar with the following license tools?

- Creative Commons License (See Q3S1)
- Open Data Commons (See Q3S2)
- Government Standard Terms of Use (Japan) (See Q3S3)

Q4. Have you ever used any of the following license tools to publish your data?

- Creative Commons License (See Q4S1)
- Open Data Commons (See Q4S2)
- Government Standard Terms of Use (Japan) (See Q4S3)

Q5. If you would like to publish your data, would you like to require the following to your users (those who use that data to publish results)?

- Noncommercial (See Q5S1)
- NoDerivs (See Q5S2)
- Secondary use prohibited (See Q5S3)
- Prohibition of use when improper use of data (See Q5S4)
- Some Penalties when improper use of data (See Q5S5)
- Fee for use (See Q5S6)
- Request to use the latest version (See Q5S7)
- Report to publish data (See Q5S8)
- Credit on the results (See Q5S9)
- Impose the same conditions when publishing results (See Q5S10)
- Nothing (freely available) (See Q5S11)

Q6. If the license is complied with, would you be willing to publish the data?

Q7. If you are using public data for your part of the research, please choose the method of presentation that you think is appropriate.

- Include source of the data information in the main text (See Q7\_1)

- Cite the source of the data in a paper (include it in the bibliography) (See Q7\_2)
- Include source of the data information in the acknowledgment (See Q7\_3)
- Add the data holder name as a co-author (See Q7\_4)
- Other (See Q7\_5)
- It is not necessary to describe the data in a paper (See Q7\_6)

Q8. Do you have any requests or concerns about using your published data for commercial activities, patents, press, literature, art, etc.?

Q9. Please select the initiatives that you believe are desirable to the use and publishing of data.

- Establishment of standard data licenses (conditions of use) (See Q9\_1)
- Development of appropriate guidelines for data licensing (See Q9\_2)
- Establishment of a data licensing consultation, support, and management department (organization) (See Q9\_3)
- Enabling a license to be specified in the data retrieval system (See Q9\_4)
- Development of data rights legislation (See Q9\_5)
- Establishment of a governing body for data licensing (external organization) (See Q9\_6)
- Other (See Q9\_7)
- Nothing in particular (See Q9\_8)

Q10. Free description (any problems or requests regarding the use or publishing of data).