

Questionnaire
A Survey on Data Citation and Reuse Practices
April, 2022

Kathleen Gregory^{1,2}, Anton Ninkov¹, Isabella Peters³, Stefanie Haustein^{1,4}

¹School of Information Studies and Scholarly Communications Lab, University of Ottawa

²University of Vienna

³ZBW - Leibniz Information Center for Economics & Kiel University

⁴Observatoire des Sciences et des Technologies (OST), Centre Interuniversitaire de Recherche sur la Science et la Technologie (CIRST), Université du Québec à Montréal

Introduction

This study investigates how researchers reuse and cite data in their work. Data reuse refers to using data (for any purpose) which others have created.

We are interested in hearing from researchers from a diversity of communities who reuse and cite data, as well as from those who do not.

The survey consists of three main sections.

- Part 1: Reusing & citing data
- Part 2: Rewarding data management
- Part 3: Demographics

This study is funded by the Alfred P. Sloan Foundation and is conducted by researchers at the [Scholarly Communications Lab](#) at the University of Ottawa, Canada as part of the [Meaningful Data Counts](#) project, a collaboration involving the nonprofit organization [DataCite](#).

By clicking on the "Next" button below to start the survey, you indicate your consent to participate in this research. You can read more about the survey and what will be done with the data on the [informed consent form](#) (this will launch a new window). We have received ethical approval for the study from the University of Ottawa Research Ethics Board.

If you need to return to a previous page as you complete the survey, please use the "Previous" button, rather than the back button in your browser.

Thank you for your participation.

Definitions

The below terms are used throughout the survey. A link to these definitions is included at the top of every page of the survey for your reference.

Data: In this survey, data are defined as entities used as evidence of phenomena for the purposes of research or scholarship [1]. Different communities define data in different ways; what serves as data for one person may not be data for another. Examples of data range from observations in spreadsheets to literature corpora to physical samples.

Secondary data: Data which others have created

Data reuse: Using data which others have created, for any purpose

Data sharing: Making your data available to others, e.g. in a data repository

Citation: Reference (to data) that is included in reference lists. Reference lists may be in formal publications (e.g. journal articles, conference proceedings, book chapters) or in other entities (e.g. blog posts, teaching syllabi, etc.).

Data management: The work involved in organizing, documenting, storing, preserving and/or sharing/publishing data in research. This work takes many forms, involving initial planning, day-to-day processes, and making data usable and available for the future.

[1] Borgman, C. L. (2015). Big data, little data, no data: Scholarship in the networked world. MIT press.

Part 1: Reusing & Citing Data

* 1. Have you ever reused data which other people have created, for any purpose?

Please select one answer.

Definitions of survey terms

Yes

No

Part 1: Reusing & Citing Data

This section is about how you reuse data which others create, so-called secondary data. We also ask questions about how you cite or otherwise reference data which you reuse.

2. Please describe the type of secondary data that you reuse (optional).

Please write your answer in the box below.

Definitions of survey terms

Part 1: Reusing & Citing Data

* 3. How would you classify the type of data that you reuse?

Please select all that apply.

Definitions of survey terms

- Quantitative data
- Qualitative data
- Both
- Other (please specify in box below)

Part 1: Reusing & Citing Data

* 4. How frequently do you reuse secondary data for the following purposes?

Please select one option for each row.

Definitions of survey terms.

	Never	Rarely	Sometimes	Often	Always
Purposes					
As the basis for a study	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
To prepare for a new project/proposal or to generate new ideas	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
To integrate with other data (e.g. to create a new dataset)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
As model, algorithm or system inputs	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
To calibrate instruments or models	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
To verify my own data	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
To identify trends, to make comparisons or to make predictions	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
To create visualizations or summaries	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
For teaching/ training	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Other (please specify in box below)

Part 1: Reusing & Citing Data

* 6. When you reuse data, how do you cite or reference them?

If you do not cite or reference data, please select never for each row.

Please select one answer for each row.

Definitions of survey terms

	Never	Rarely	Sometimes	Often	Always
Including a citation to data in reference lists	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Including a citation to a related paper in reference lists	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Mentioning data in a footnote	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Mentioning data in body of text	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Mentioning data in acknowledgements, appendix, etc.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Mentioning data in captions, figures or tables	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Other (please specify in box below)

Part 1: Reusing & Citing Data

* 7. If **another person** were to reuse data which **you** have created, would you want them to cite/reference your data?

Please select one answer.

Yes

No

Part 1: Reusing & Citing Data

This page contains two related questions about your preferences regarding how other people cite or reference your data.

* 8. What would you prefer that other people cite/reference when they use **your** data?

Please select all that apply.

Definitions of survey terms

- The data themselves (e.g. a particular dataset or record)
- The source of the data (e.g. a data repository or database)
- A collection of related data (e.g. a series of datasets)
- An article or publication analyzing the data
- A data paper (A data paper is a short article with in-depth descriptions about the particular data).
- Other (please specify in box below)

* 9. How would you prefer other people to cite/reference **your** data?

Please select all that apply.

- Including a citation to data in reference lists
- Including a citation to a related paper in reference lists
- Mentioning data in a footnote
- Mentioning data in body of text
- Mentioning data in acknowledgements, appendix, etc.
- Mentioning data in captions, figures or tables
- Other (please specify in box below)

Part 1: Reusing & Citing Data

This section asks further questions about your motivations for citing (or not citing) data. Citations are references to data, in reference lists.

* 10. Do you cite secondary data (by including a data citation in a reference list) in any way for any purpose?

Please select one answer.

Definitions of survey terms

- Yes
- Sometimes
- No

If you selected "No," please explain why you do **not** cite data.

Part 1: Reusing & Citing Data

* 11. Why do you cite secondary data?

Please select all that apply.

Definitions of survey terms

- | | |
|--|---|
| <input type="checkbox"/> As a way of showing intellectual debt to the data creator/data provider | <input type="checkbox"/> As a way of rewarding the data creators/data providers |
| <input type="checkbox"/> As a way of supporting the validity of your claims | <input type="checkbox"/> As a way of indicating that you have used the data in some way |
| <input type="checkbox"/> As a way of helping others to locate and access the data you used | <input type="checkbox"/> Because you were advised to cite data (e.g. by a journal editor) |
| <input type="checkbox"/> Other (please specify in box below) | |

Part 1: Reusing & Citing Data

* 12. Do you ever cite data for the following reasons?

Please select all that apply.

Definitions of survey terms

- | | |
|--|---|
| <input type="checkbox"/> To correct your own data (you cite your own data) | <input type="checkbox"/> To bring recognition to your own data (you cite your own data) |
| <input type="checkbox"/> To build on or use data you have created (you cite your own data) | <input type="checkbox"/> None of the above |
| <input type="checkbox"/> To criticize or correct the data of others | |

Part 1: Reusing & Citing Data

* 13. What are your reasons for not reusing data created by other people?

Please select all that apply.

Definitions of survey terms

- | | |
|--|--|
| <input type="checkbox"/> Reusing data is not relevant to your research methods | <input type="checkbox"/> You do not trust data that others create |
| <input type="checkbox"/> Reusing data is not normal in your research community | <input type="checkbox"/> You cannot find the data you need |
| <input type="checkbox"/> There are no available relevant data for your research topic | <input type="checkbox"/> You do not know how to give credit for data created by other people |
| <input type="checkbox"/> You get more credit for your work if you create your own data | <input type="checkbox"/> You did not know that you could reuse data created by other people |
| <input type="checkbox"/> Other (please specify in the box below) | |

Part 1: Reusing & Citing Data

* 14. Please describe the type of data that you use in your work.

Please write your answer in the box below.

Definitions of survey terms

Part 1: Reusing & Citing Data

* 15. If **another person** were to reuse data which **you** created, would you want them to cite/reference your data?

Please select one answer.

Yes

No

Part 1: Reusing & Citing Data

This page contains two related questions about your preferences regarding how other people cite/reference your data.

* 16. What would you prefer that other people cite/reference when they use your data?

Please select all that apply.

Definitions of survey terms

- The data themselves (e.g. a particular dataset or record)
- The source of the data (e.g. a data repository or database)
- A collection of related data (e.g. a series of datasets)
- An article or publication analyzing the data
- A data paper (A data paper is a short article with in-depth descriptions about the particular data).
- Other (please specify in box below)

* 17. How would you prefer other people to cite/reference **your** data?

Please select all that apply.

- Including a citation to data in reference lists
- Including a citation to a related paper in reference lists
- Mentioning data in a footnote
- Mentioning data in body of text
- Mentioning data in acknowledgements, appendix, etc.
- Mentioning data in captions, figures or tables
- Other (please specify in box below)

Part 1: Reusing & Citing Data

* 18. Are you aware of and do you use recommendations/standards for **citing data** from the following bodies?
Please select one option in each row.

Definitions of survey terms

	I am aware of and use this recommendation.	I am aware of and do not use this recommendation	I am unaware of this recommendation.
DataCite	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Data repositories	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Scientific societies (e.g. the American Geophysical Union)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Journals or publishers	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Citation style guides (e.g. APA, MLA, etc.)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Part 2: Rewarding Data Management

This section asks questions about rewarding and documenting data management, data sharing and data reuse.

* 19. How important are the following to you?

Please select one option for each statement.

Definitions of survey terms

	Extremely unimportant	Unimportant	Neither important nor unimportant	Important	Extremely important
Having data be recognized as a standalone research output, separate from publications	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Having the creation of good data documentation and workflows be rewarded	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Part 2: Rewarding Data Management

This page contains two questions which are related to each other.

* 20. Have you ever shared your own research data?

Please select one answer.

Definitions of survey terms

Yes

No

* 21. Have you ever used **your own** data multiple times (e.g. for new studies, or to produce multiple publications?)

Please select one answer.

Definitions of survey terms

Yes

No

Part 2: Rewarding Data Management

This page contains two questions which are related to each other.

* 23. How important would it be for you to know the following information about **others'** data which you may potentially reuse?

Please select one option for each row.

	Extremely unimportant	Unimportant	Neither important or unimportant	Important	Extremely important
The number of citations the data have received	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The number of times the data have been downloaded	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The number of times the data have been viewed	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Information about where the data were used (e.g., in which country, at which institution)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Descriptions or a narrative providing details about how the data were used	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Information about who has used the data	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
If the data have received recognition outside the scholarly system, e.g. in newspapers	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

* 24. How important would it be for you to know the following information about **your** data?

Please select one option for each row.

Definitions of survey terms

	Extremely unimportant	Unimportant	Neither important or unimportant	Important	Extremely important
The number of citations your data have received	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The number of times your data have been downloaded	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The number of times your data have been viewed	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Information about where your data were used (e.g., in which country, at which institution)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Descriptions or a narrative providing details about how your data were used	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Information about who has used your data	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
If your data have received recognition outside the scholarly system, e.g. in newspapers	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Part 3: Demographics

You are nearly at the end of the survey. Below are some questions to help us contextualize your answers.

* 25. How many years of professional experience do you have in your field?

Please select one answer.

0-5

6-15

16-30

31+

Part 3: Demographics

* 26. With which disciplinary domain do you most identify ?

Please select one discipline which best describes your expertise.

If desired, please add keywords to further describe your disciplinary expertise below.

Part 3: Demographics

* 28. What type of organization do you work for?

Please select one answer.

- University or college
- Research institution
- Other (please specify in box below)
- Government agency
- Corporate

Part 3: Demographics

* 29. In what country do you currently work?

Part 3: Demographics

* 30. How do you identify your gender:

Please select one answer.

- Woman
- Man
- Non-binary
- Prefer not to say
- Option not listed

Part 4: Conclusion

In this section you can add final comments and indicate if you would like to be contacted in the future.

31. Final comments: Do you have anything else that you would like to share with us?

Please write your comments in the box below.

Part 4: Conclusion

32. May we contact you in the future?

Please select one answer per row.

	Yes	No
I would like to be informed about the results of this survey.	<input type="radio"/>	<input type="radio"/>
I would be willing to participate in future research in this area.	<input type="radio"/>	<input type="radio"/>

33. Please enter your email address below if you would like to be contacted.
If you would not like to be contacted please proceed to the next page.

Email addresses will only be used to contact you for the previously mentioned purposes. Your email address will be stored in a location separate from your survey responses.

**You have reached the end of the survey.
Thank you very much for your time and input.**

Please click "Done" to complete the survey.

The Meaningful Data Counts Team

**Dr. Stefanie Haustein, Dr. Kathleen Gregory, Dr. Anton Ninkov, & Dr. Isabella Peters
stefanie.haustein@uottawa.ca**