Plan Overview

A Data Management Plan created using DMPonline

Title: Collecting Viewpoint Annotations for Online Content From Crowd Workers

Creator: Tim Draws

Affiliation: Delft University of Technology

Template: TU Delft Data Management Plan template (2021)

Project abstract:

The aim of this project is to have crowd workers annotate the viewpoints on different debated topics expressed in pieces of online content (e.g., tweets, forum entries, news articles). We also ask crowd workers for their own opinion and knowledge on these debated topics.

ID: 83932

Start date: 20-09-2021

End date: 15-10-2021

Last modified: 08-09-2021

Copyright information:

The above plan creator(s) have agreed that others may use as much of the text of this plan as they would like in their own plans, and customise it as necessary. You do not need to credit the creator(s) as the source of the language used, but using any of the plan's text does not imply that the creator(s) endorse, or have any relationship to, your project or proposal

Collecting Viewpoint Annotations for Online Content From Crowd Workers

0. Administrative questions

1. Name of data management support staff consulted during the preparation of this plan.

2. Date of consultation with support staff.

Question not answered.

I. Data description and collection or re-use of existing data

3. Provide a general description of the type of data you will be working with, including any re-used data:

Type of data	File format(s)	How will data be collected (for re-used data: source and terms of use)?	Purpose of processing	Storage location	Who will have access to the data
(Anonymised) personal opinions and knowledge of crowd workers on debated questions	csv	crowdsourcing task (Amazon Mturk)	this is part of what we study	project drive	me and the collaborators on this project
Crowd workers' viewpoint annotations for online content	see above	see above	see above	see above	see above

4. How much data storage will you require during the project lifetime?

- < 250 GB

II. Documentation and data quality

5. What documentation will accompany data?

- Methodology of data collection
- README file or other documentation explaining how data is organised
- Data dictionary explaining the variables used
- Data will be deposited in a data repository at the end of the project (see section V) and data discoverability and re-usability will be ensured by adhering to the repository’s metadata standards

III. Storage and backup during research process

6. Where will the data (and code, if applicable) be stored and backed-up during the project lifetime?

- Project Storage at TU Delft
- SURFdrive

- Another storage system - please explain below, including provided security measures

Anything we upload to surfdrive or other (public) platforms (e.g., Open Science Framework) will be anonymised.

IV. Legal and ethical requirements, codes of conduct

7. Does your research involve human subjects or 3rd party datasets collected from human participants?

- Yes

8A. Will you work with personal data? (information about an identified or identifiable natural person)

If you are not sure which option to select, ask your Faculty Data Steward for advice. You can also check with the privacy website or contact the privacy team: privacy-tud@tudelft.nl

- Yes

8B. Will you work with any types of confidential or classified data or code as listed below? (tick all that apply)

If you are not sure which option to select, ask your Faculty Data Steward for advice.

- No, I will not work with any confidential or classified data/code

9. How will ownership of the data and intellectual property rights to the data be managed?

For projects involving commercially-sensitive research or research involving third parties, seek advice of your Faculty Contract Manager when answering this question. If this is not the case, you can use the example below.

The datasets underlying the published papers will be publicly released following the TU Delft Research Data Framework Policy. During the active phase of research, the project leader from TU Delft will oversee the access rights to data (and other outputs), as well as any requests for access from external parties. They will be released publicly no later than at the time of publication of corresponding research papers.

10. Which personal data will you process? Tick all that apply

- Gender, date of birth and/or age
- Special categories of personal data (specify which): race, ethnicity, criminal offence data, political beliefs, union membership, religion, sex life, health data, biometric or genetic data
- Other types of personal data - please explain below

When collecting data via Amazon Mturk, the Mturk IDs of workers are automatically transmitted to us. However, we do not need them so we will anonymise the data as soon as possible.

11. Please list the categories of data subjects

- crowd workers (on Amazon Mechanical Turk)

12. Will you be sharing personal data with individuals/organisations outside of the EEA (European Economic Area)?

- No

15. What is the legal ground for personal data processing?

- Informed consent

16. Please describe the informed consent procedure you will follow:

I write the informed consent myself, together with collaborators.

17. Where will you store the signed consent forms?

- Same storage solutions as explained in question 6

18. Does the processing of the personal data result in a high risk to the data subjects?

If the processing of the personal data results in a high risk to the data subjects, it is required to perform a Data Protection Impact Assessment (DPIA). In order to determine if there is a high risk for the data subjects, please check if any of the options below that are applicable to the processing of the personal data during your research (check all that apply).
If two or more of the options listed below apply, you will have to complete the DPIA. Please get in touch with the privacy team: privacy-tud@tudelft.nl to receive support with DPIA.
If only one of the options listed below applies, your project might need a DPIA. Please get in touch with the privacy team: privacy-tud@tudelft.nl to get advice as to whether DPIA is necessary.
If you have any additional comments, please add them in the box below.

- None of the above applies

22. What will happen with personal research data after the end of the research project?

- Personal data will be shared with others - please explain which personal data will be shared, with whom, how and whether you have specified this in the informed consent form

We will not share any Mturk IDs or other data that could trace back to the individual crowd worker. We may publicly share all other data we collect in an anonymised format; together with other research output (e.g., paper).

23. How long will (pseudonymised) personal data be stored for?

- 10 years or more, in accordance with the TU Delft Research Data Framework Policy

24. What is the purpose of sharing personal data?

- For research purposes, which are in-line with the original research purpose for which data have been collected

25. Will your study participants be asked for their consent for data sharing?

- Yes, in consent form - please explain below what you will do with data from participants who did not consent to data sharing

We explain in the informed consent that data may be shared publicly in an anonymised format.

V. Data sharing and long-term preservation

27. Apart from personal data mentioned in question 22, will any other data be publicly shared?

- All other non-personal data (and code) produced in the project

29. How will you share research data (and code), including the one mentioned in question 22?

- I will upload the data to another data repository (please provide details below)

Open Science Framework

31. When will the data (or code) be shared?

- As soon as corresponding results (papers, theses, reports) are published

VI. Data management responsibilities and resources

33. Is TU Delft the lead institution for this project?

- Yes, leading the collaboration

34. If you leave TU Delft (or are unavailable), who is going to be responsible for the data resulting from this project?

Prof. Geert-Jan Houben (Web Information Systems Group)

35. What resources (for example financial and time) will be dedicated to data management and ensuring that data will be FAIR (Findable, Accessible, Interoperable, Re-usable)?

financial and time