Plan Overview

A Data Management Plan created using DMPonline

Title: Does multi-party computation (MPC) enhance control in data sharing through data marketplaces? An experimental study

Creator: Wirawan Agahari

Principal Investigator: Wirawan Agahari

Data Manager: Wirawan Agahari

Project Administrator: Wirawan Agahari

Affiliation: Delft University of Technology

Template: TU Delft Data Management Plan template (2021)

Project abstract:

Although data sharing via data marketplaces poses numerous benefits in the data economy, individuals are reluctant to share data with third parties due to lack of trust, fear of losing control over data, and privacy concerns. Multi-party computation (MPC) is a cryptographic technique that enables multiple parties to jointly analyze data while retaining the secrecy of the data. While MPC could overcome risks that emerge due to platformization of data sharing, the meaning of MPC in data marketplaces setting and its influence on the willingness to share data is not yet researched. This research aims to investigate the impact of MPC in enhancing individuals' perception of control over data and their willingness to share data through data marketplaces. To do so, we build upon information privacy theory to construct a theoretical model on the impact of MPC on perceived control over data and individuals’ willingness to share data in data marketplaces setting. This theoretical model will be tested via pre-test-post-test between-subjects experimental study using a mockup of MPC-enabled data marketplaces. We expect our findings to serve as a foundation for future research in the emerging phenomenon of platformization of data sharing via data marketplaces and the key role of MPC in enabling the data economy.

ID: 80843

Start date: 01-03-2021

End date: 28-02-2023

Last modified: 24-08-2021

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

Does multi-party computation (MPC) enhance control in data sharing through data marketplaces? An experimental study

0. Administrative questions

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

My PhD promotor (Mark de Reuver) and co-promotor (Tobias Fiebig), have reviewed this DMP on the 6th of July, 2021.
My faculty data steward, Nicolas Dintzner, has reviewed this DMP on the 12th of July, 2021.

2. Date of consultation with support staff.

2021-07-12

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
Anonymized demographic data (age, gender, occupation, education level, car ownership, country of residence)	.csv files	online survey via Qualtrics	To describe demographic profiles of participants and make appropriate correlations	Qualtrics server and SURF Drive	The PI (Wirawan Agahari), supervisors (Mark de Reuver, Tobias Fiebig), and research assistant (Iris van der Wel)
Anonymized data on perceived control, trust, perceived risk, privacy concerns, and willingness to share	.csv files	online survey via Qualtrics	To understand individuals' perception of variables under study based on different treatments (no MPC, MPC, and made-up privacy technology)	Qualtrics server and SURF Drive	The PI (Wirawan Agahari), supervisors (Mark de Reuver, Tobias Fiebig), and research assistant (Iris van der Wel)
Anonymized data on Westin's Privacy Index	.csv files	online survey via Qualtrics	To determine the privacy profile of each respondent and make appropriate correlations	Qualtrics server and SURF Drive	The PI (Wirawan Agahari), supervisors (Mark de Reuver, Tobias Fiebig), and research assistant (Iris van der Wel)

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

All datasets and metadata will follow the DataCite standard because they will be published in 4TU Research Data Repository.

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?

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

Qualtrics server (will be removed once the project is finished).

IV. Legal and ethical requirements, codes of conduct

7. Does your research involve human subjects?

- 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

11. Please list the categories of data subjects

The data subject for this study is UK residents who currently live in the UK. I will use the online participant recruitment service called Prolific to identify and recruit relevant data subjects and ensure the representativeness of the sample to the UK population. My supervisor is already got a confirmation from the legal department of TU Delft that there is no issue regarding the GDPR for the use of Prolific.

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:

All study participants will be asked for their consent (i.e., clicking "I consent to participate in this study) for taking part in the study and for data processing before the start of the survey. Those who do not consent should not even partake in the study and will immediately be redirected to the end of the survey and stating that they cannot participate. Participants also have a right to withdraw from the study at any point.

Dear participants,

Thank you for making time to take part in this survey. Your contribution is greatly appreciated!

You are being invited to participate in a research study titled **"Does multi-party computation (MPC) enhance control in data sharing through data marketplaces? An experimental study."** This study is being done by Wirawan Agahari, a Ph.D. researcher at Delft University of Technology (TU Delft), supervised by Dr. Ir. Mark de Reuver and Dr.-Ing. Tobias Fiebig.

The purpose of this research study is **to investigate the impact of a privacy-enhancing technology called Multi-Party Computation (MPC) on individuals' control over data and willingness to share data in data marketplaces.** This study will take you **approximately 30 minutes** to complete. Your answer will remain anonymous, cannot be traced back to you, and will only be used for research purposes. **Your participation in this study is entirely voluntary and you can withdraw at any time.**

We believe there are no known risks associated with this research study; however, as with any online-related activity, the risk of a breach is always possible. To the best of our ability, your answers in this study will remain confidential. We will minimize any risks by **only storing data at remote, protected storage at TU Delft, only accessible by project members as well as abstaining from both distributing data to others or retrieving it on personal devices.**

An anonymized, non-reducible version of this dataset will be publicly available through 4TU Research Data Repository. Before publication, any personal data will be dropped.

For any further inquiries, please refer to:

Wirawan Agahari
Ph.D. researcher
Delft University of Technology
w.agahari@tudelft.nl
Please check the first box to give permission to process your data for this research:

- I acknowledge that I have read and understood this introduction, and **I hereby give consent** that my survey data will be processed for this research.
- **I do not consent**, I do not wish to participate in this study.

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?

- Anonymised or aggregated data will be shared with others
- Personal research data will be destroyed after the end of the research project

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?

- Other - please explain below

No personal data will be collected throughout this survey, and therefore there will be no personal data shared with other parties.

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 will ask participants who did not consent to data sharing (i.e., clicking "I do not consent, I do not wish to participate in this study") to not participate in the survey. This is done by redirecting them to the end of the survey.

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) underlying published articles / reports / theses

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

- All anonymised or aggregated data, and/or all other non-personal data will be uploaded to 4TU.ResearchData with public access

30. How much of your data will be shared in a research data repository?

- < 100 GB

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

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

32. Under what licence will be the data/code released?

- CC BY

VI. Data management responsibilities and resources

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

- Yes, leading the collaboration

This PhD research is part of Work Package 2 (WP2) at the EU H2020 project Safe-DEED, in which TU Delft is the lead.

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

Mark de Reuver, Head of ICT section, ESS Department, TPM Faculty (g.a.dereuver@tudelft.nl)
Mark is also my PhD promotor and my daily supervisor.

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)?

- 4TU.ResearchData is able to archive 1TB of data per researcher per year free of charge for all TU Delft researchers. We do not expect to exceed this and therefore there are no additional costs of long-term preservation.
- Possible publishing costs for (1) data publication in data repositories other than 4TU.ResearchData (where TU Delft researchers

can publish free of charge for up to 1TB of data per researcher per year); and (2) publication of papers about datasets or software in dedicated journals.