Plan Overview

A Data Management Plan created using DMPonline

Title: Understanding how people explain the behaviour of autonomous vehicles

Creator:Balint Gyevnar

Principal Investigator: Balint Gyevnar

Data Manager: Balint Gyevnar

Affiliation: University of Edinburgh

Funder: UKRI Future Leaders Fellowships

Template: UKRI Future Leaders Fellowships Template for a Data Management Plan

ORCID iD: 0000-0003-0630-4315

Project abstract:

This project aims to understand the kinds of explanations people tend to give, and prefer to receive, regarding the driving behaviour of autonomous vehicles (self-driving cars) in complex simulated scenarios. We are interested in how people perceive different kinds of natural language explanations and how these different explanations affect people's understanding of how vehicles will behave. As part of this project, we aim to create and publish an anonymised dataset of natural language explanations about interesting and challenging driving scenarios and hope to gain a better understanding of how and when people require explanations about autonomous driving.

ID: 122065

Last modified: 12-04-2023

Grant number / URL: EP/S022481/1

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

Understanding how people explain the behaviour of autonomous vehicles

0. Proposal name

0. Enter the proposal name

Understanding how people explain the behaviour of autonomous vehicles

1. Description of the data

1.1 Type of study

Online survey for understanding the kinds of explanations people tend to give, and prefer to receive, regarding the driving behaviour of autonomous vehicles (self-driving cars) in complex simulated scenarios.

1.2 Types of data

All data is collected as part of an online survey.

- Quantitative:
  - Demographic data
  - Driving experience-related data
  - Preference judgments about different explanations of autonomous vehicles
- Qualitative:
  - Free text explanations about autonomous vehicles

1.3 Format and scale of the data

Response data from human participants in Excel amounting to a few MBs in size.

2. Data collection / generation

2.1 Methodologies for data collection / generation

Randomised online survey with human participants using a pre-determined set of driving scenarios

2.2 Data quality and standards

The uniform quality of the data is guaranteed by the use of an online survey platform such as Qualitrics. Participants are recruited via the online crowdsourcing platform Prolific and paid a (pro rata) living wage to guarantee that high-quality data is collected. Additionally, attention and eligibility checks will serve to filter out intelligible or inattentive participants.

3. Data management, documentation and curation

3.1 Managing, storing and curating data

All data will be processed in accordance with Data Protection Law. Any personal or sensitive information collected about participants will be kept strictly confidential. Data will be referred to by a unique participant number rather than by name. All electronic data containing any personal or sensitive information will be stored on a password-protected encrypted computer, on the School of Informatics' secure file servers, or on the University's secure encrypted cloud storage services (DataShare, ownCloud, or Sharepoint).

3.2 Metadata standards and data documentation

The data will be accompanied by a README.txt file that contains common metadata including title and abstract, keywords, data creators, funder, data type (e.g. image, text, data set), publisher (e.g. UoE, School of Chemistry), grant id, and related documents (e.g. journal articles).

The file will contain a layout of the file structure of the data including a description of each file, their extensions, a summary of their contents, and how they can be used.

The same file will also contain metadata about the fields of the dataset including a brief summary of the field, its data type and expected values.

3.3 Data preservation strategy and standards

The data will be preserved on DataShare for a minimum of 2 years with the accompanying documentation and metadata contained in the README.txt file.

4. Data security and confidentiality of potentially disclosive information

4.1 Formal information/data security standards

All data will be processed in accordance with Data Protection Law. Any personal or sensitive information collected about participants will be kept strictly confidential. Data will be referred to by a unique participant number rather than by name. All electronic data containing any personal or sensitive information will be stored on a password-protected encrypted computer, on the School of Informatics' secure file servers, or on the University's secure encrypted cloud storage services (DataShare, ownCloud, or Sharepoint).

4.2 Main risks to data security

The main risk to the data stored on an online encrypted data-sharing platform, such as DataShare, is an external breach of the sharing platform's servers. While this is out of the researchers' control, we will anonymise all personal or sensitive data when publishing it online and redact all personal or sensitive information from free text responses.

5. Data sharing and access

5.1 Suitability for sharing

Yes, it is suitable for sharing. Quotes or key findings will be anonymized: We will remove any information that could, in our assessment, allow anyone to identify participants. Any personal or sensitive information in free text responses will be removed and we will not publish demographics-related data fields.

5.2 Discovery by potential users of the research/innovation data

Potential users can access the data via a publicly available link to the data-sharing platform, which will be published as part of the report that will follow the data collection effort. This link will be available to anyone with an internet connection and access to the link.

5.3 Governance of access

The data will be publicly available without the restriction of access in a data-sharing platform.

5.4 The study team’s exclusive use of the data

The study team maintains a maximum of 6 months of exclusive access to the data before releasing it online.

5.5 Restrictions or delays to sharing, with planned actions to limit such restrictions

A consent procedure regarding data-sharing and consent withdrawal will be included as part of the data collection such that all consenting participants' data can be included in the publicly available dataset without future restrictions. If consent is withdrawn from a participant then the public dataset will be promptly updated to remove their data.

5.6 Regulation of responsibilities of users

The data may be accessed, used, and distributed only in its original form by any external users.

6. Responsibilities

6. Responsibilities

The PI takes all responsibility for data management, metadata creation, data security, and quality assurance of the data.

7. Relevant policies

7. Relevant institutional, departmental or study policies on data sharing and data security

Policy	URL or Reference
Data Management Policy & Procedures	https://www.ed.ac.uk/information-services/about/policies-and-regulations/research-data-policy)
Data Security Policy	https://www.ed.ac.uk/data-protection/data-protection-policy
Data Sharing Policy	https://www.ed.ac.uk/records-management/guidance/requests
Institutional Information Policy	https://www.ed.ac.uk/data-protection/data-protection-policy
Other	
Other	

8. Author and contact details

8. Author of this Data Management Plan (Name) and, if different to that of the Principal Investigator, their telephone & email contact details

Balint Gyevnar, balint.gyevnar@ed.ac.uk