A picture containing object, clock, screen, laptop

Description automatically generated

**Data Management Plan Hot Politics Lab, version 1.1**

Gijs Schumacher

**Data creation and collection**

We typically collect data through 1) experiments, 2) surveys, 3) analysis of secondary survey material, and 4) public speeches and videos.

1. Physiological and attitudinal data collected in lab experiments

We will collect physiological measures such as skin conductance, heartbeat and facial electromyography. To collect this we use the vsrrp98 software, which was developed in-house. The data is stored .vrdf files, and is typically quite large because we sample data per millisecond in an experiment. A data collection of 150 respondents typically results in 1GB of data. We follow set procedures in vsrrp98 to clean the data.

Attitudinal data in the lab will be collected through Qualtrics and Presentation. The latter outputs log files. Data can be exported from Qualtrics as a .csv file. will not export IP-addresses from Qualtrics. Both types of files will be small in size (<1 MB).

All the data we collect will be owned by the University and no one else. Other parties can use the data we publish onto the public domain.

1. Attitudinal data collected through survey experiments

We use different survey companies to collect data. We always have ownership of the data and will ask for survey responses in .csv format (again, small size <1 MB). We will not export personal identifiable data such as IP-addresses.

1. Secondary survey data

We analyze secondary data such as election surveys, household panels and data collected by other researchers. We follow the guidelines and procedures of the owner of the data when we store and share this data.

1. Transcripts, audio and video materials of politicians.

We use official transcripts, audio and video materials of politicians published in official repositories such as those of national parliaments (in the Netherlands, for example: [https://debatgemist.tweedekamer.nl](https://debatgemist.tweedekamer.nl/)). These sources are public and can be used for research ends.

**Data management, documentation and curation**

Here we will use the same protocol regardless of the type of data. We will distinguish between storing data that is actively used for research and published data.

Data in use for research

All the data that is collected and analyzed will initially be saved on Surfdrive (a safe repository designed for Dutch academia). We will store the raw data in DATA folders. In PAPER folders we will store scripts that amend the data, and save the amended data as new files in the respective folder. For clarification, the folder structure on each laptop should be like this:

For raw data: C:\Users\uva-id\surfdrive\data\data collection 1\

For analyses: C:\Users\uva-id\surfdrive\papers\paper 1\

We will set up standardized scripts to facilitate this workflow.

Published research

To make the workflow open in principle all materials will be uploaded to the Open Science Framework (scripts, datasets (without personally identifiable data), experimental or survey materials (if applicable), pre-analysis plan and pre-print). This way we can guarantee access. There are two exceptions to this principle. First, if user agreements of secondary data stipulate that the data cannot be made open access, we do not publish it (click [here](https://dataverse.harvard.edu/file.xhtml?persistentId=doi:10.7910/DVN/HJSQNL/OAIEZO&version=1.0) for an example). We do add instructions to our scripts how to obtain the data. Second, if raw data exceeds 5GB it cannot be uploaded to OSF. We make it available through a surfdrive link. We will add the following metadata in README files: title, creator(s), description and date in the OSF folder. We will keep a copy of the OSF folder on surfdrive.

Note that there is no personally identifiable data in the data we share.

For the text projects the scripts to replicate the entire paper will also be uploaded to Github.

We are not collecting any non-digital data.

**Data security**

For our lab studies we need the names, addresses and bank numbers of our participants so that they can get paid. We administer this on a separate sheet than the data collection. So no link can be made between the personally identifiable information and the attitudes and responses of the participants. To pay participants we email this sheet to the relevant administration (secretariat AISSR or ASCOR, depending on who pays for the experiment). This email needs to be send using filesender from surfdrive. This way the data is encrypted and the personal data is transferred in a secure way. After this, we destroy the personal data.

In the remainder of the data collection / creation no personally identifiable information is collected, therefore it is not necessary to apply specific security measures.

At all times we will work on university managed laptops to ensure the highest possible level of security against data leaks.

After publication research data will no longer be stored on individual laptops, to limit the possibility of data leaks in case laptops are stolen.

**Data archival and preservation**

All data collected in the project will be maintained. Data will be uploaded to OSF, and copies will be kept on surfdrive (in the cloud). The data and thereby long-term access and availability is guaranteed.

**Data publication and access**

In principle all data (from raw data to the results published in a paper) will be made available through OSF (see section Data management for more detail). This way the entire research process could be replicated by third parties. On OSF we will also upload the README files so that relevant metadata is available too. The data will be under embargo until the online publication of the associated paper.

**Roles, responsibilities and resourcing**

The lead author of a paper will bear the responsibility of the implementation of this plan. In case the lead author is a PhD student, the promotor of the PhD students bears the responsibility of the implementation.

The internal director of the Hot Politics Lab will communicate the data management plan to the team members. The internal director of the Hot Politics Lab will check the implementation of the data management plan annually and will consider updating it.