

# MY PLAN (FWO DMP)

## Data Management Plan: InFormMe: Influencers' (Mis-)Information on Social Media and Adolescents' Political Participation

### ADMIN DETAILS

**Project Name:** InFormMe: Influencers' (Mis-)Information on Social Media and Adolescents' Political Participation

**Principal Investigator / Researcher:** Desiree Schmuck

**Institution:** KU Leuven

### 1. GENERAL INFORMATION

**Name applicant**

*Desiree Schmuck*

**FWO Project Number & Title**

*G0B7822N; InFormMe: Influencers' (Mis-)Information on Social Media and Adolescents' Political Participation*

**Affiliation**

*KU Leuven, School for Mass Communication Research*

### 2. DATA DESCRIPTION

**Will you generate/collect new data and/or make use of existing data?**

*New data will be generated in all work packages of the project.*

**Describe in detail the origin, type and format of the data (per dataset) and its (estimated) volume. This may be easiest in a table (see example) or as a data flow and per WP or objective of the project. If you reuse existing data, specify the source of these data. Distinguish data types (the kind of content) from data formats (the technical format).**

Type of data	Format	Volume	Work package	Procedure
Survey data, images (screenshots)	.sav, .jpg	5 GB	WP1	Via a mobile application, participants (adolescents) in the citizen science study provide information about influencers on social media who post political (mis-)information and take screenshots of influencers' posts. They will be actively involved in the research by also being asked to categorize the content that they

				upload in the mobile application. This tool, which will be most likely an application provided by the experienced software developer ilumivu, can be installed on citizen scientists' smartphones, if they decide to participate in the project. The involved citizen scientists can only upload data but have no way of accessing data uploaded by others. The company's privacy policy ensures safe handling of data and the application called mEMA has been used by numerous universities for academic mobile experience sampling research. The uploaded data will be stored on KU Leuven's central network drive that can only be accessed by the PI and the primary project team.
Recorded focus group discussions (audio)	.mp3	1 GB	WP1	An "advisory board" of 20 adolescent participants will help plan the research by participating in focus group discussions, which will be recorded (audio only) to plan and reflect the research.
Content analytical data (numerical)	.sav	100 MB	WP1	Using a codebook implemented into Qualtrics, coders will analyze N = 2000 posts/stories for the presence of political (mis-)information of those influencers (N= 50) who were most frequently nominated in the citizen science study.
Pilot survey and longitudinal survey data (numerical)	.sav	100-500 MB	WP2	In online surveys (via Qualtrics) filled out during class in computer rooms, adolescents will be asked about their exposure and relationship to influencers as well as their perceptions of 'political' influencers. The surveys will also measure their political opinions and attitudes (e.g., political engagement, political participation). Adolescents will also nominate close peers via a unique code, who also fill out surveys via Qualtrics. This procedure allows linking adolescents to their peers without linking personal data like names or email addresses to the survey responses. Parents of adolescents will be asked to respond to online surveys (Qualtrics) about how they mediate their children's social media use. Parents' and children's data will again be linked via the same unique code.

### 3. LEGAL AND ETHICAL ISSUES

**Will you use personal data? If so, shortly describe the kind of personal data you will use. Add the reference to your file in KU Leuven's Register of Data Processing for Research and Public Service Purposes (PRET application). Be aware that registering the fact that you process personal data is a legal obligation.**

*In both WPs, self-report data related to adolescents' exposure and relationship to influencers and their political content will be collected. Moreover, data about their political views and attitudes (e.g., political engagement) will be collected. Data is collected via a mobile application (WP1) and the survey application Qualtrics (WP2). An ethical application has been submitted to PRET (G-2022-5237) to receive ethical approval for the planned work packages and the collection of personal data. The application was reviewed and received the decision "Major revision." The application is currently revised by the authors.*

**Are there any ethical issues concerning the creation and/or use of the data (e.g. experiments on humans or animals, dual use)? If so, add the reference to the formal approval by the relevant ethical review committee(s)**

*No data collection is planned which is based on experiments.*

**Does your work possibly result in research data with potential for tech transfer and valorisation? Will IP restrictions be claimed for the data you created? If so, for what data and which restrictions will be asserted?**

*No.*

**Do existing 3rd party agreements restrict dissemination or exploitation of the data you (re)use? If so, to what data do they relate and what restrictions are in place?**

*No.*

### 4. DOCUMENTATION AND METADATA

**What documentation will be provided to enable reuse of the data collected/generated in this project?**

*For each WP study, we will create a separate folder that contains the following information:*

- 1. Read me.docx: In this word-document, we will provide an overview of which researchers were involved in the collection of the data (e.g., PhD students, Master students, contact persons in schools, responsible teachers), a short overview of the study course and protocol, which questionnaires we administered and their labels, and short written information on the data cleaning process and steps.*
- 2. Email addresses.xlsx: In this excel-document, we will store email addresses that are assessed in WP2 to 1) send parents the parent survey and (re-)contact them for the second and third survey waves, 2) provide parent and adolescent participants with the aggregated results of the research*

and 3) send parent and adolescent participants the vouchers (incentives) for the participation. This identifying information cannot be linked with the responses in the survey, because it is assessed via a separate form in Qualtrics. This document will be deleted as soon as not needed anymore and only a completely anonymized dataset will be kept.

3. Folder with all the study documents: Ethical application and approval (G-2022-5237; KU Leuven), examples of parental and informed consent files, as well as the information brochures we gave participants. The PDF of all questionnaires (both baseline and ESM) will be included.

4. Printed parental consent files: Parental consent forms will be collected in written form and will contain identifying information by parents and their children. These files are printed paper documents and will be stored in a locked closet in the principal investigator's office at KU Leuven. The files will be discarded in line with the policy at KU Leuven when not needed anymore.

5. Pre-processing documents: The raw data (anonymized csv files) of WP 1 and WP 2.

6. Documents for data analysis: The processed anonymized data (.R or .sps format). A reproducible syntax/code for the data analysis (either .R or .sps format).

7. Open Science: Documents such as data scripts and datasets with survey responses (but without identifying information) will also be made open access via the Open Science Framework (osf.io), in order to allow reproduction of the data. The folder will never contain sensitive (identifiable) information about participants such as names or email addresses.

**Will a metadata standard be used? If so, describe in detail which standard will be used. If no, state in detail which metadata will be created to make the data easy/easier to find and reuse.**

*We will use metadata from our survey studies including date of participation, duration of survey response, browser type, user language, progress in the survey, question text.*

## **5. DATA STORAGE AND BACKUP DURING THE FWO PROJECT**

**Where will the data be stored?**

*All data will be stored on secure KU Leuven servers (I-drive or J-drive). Data from WP1 and WP2 will be stored on the encrypted network J-drive, and is only accessible to involved researchers using their employee ID and password. Restricted and sensitive data (i.e., email addresses) will be stored separately on an encrypted password-protected personal I-drive of the KU Leuven for the duration of this project, and can only be accessed by the main researchers of this project. Both drives are backed-up automatically to prevent loss of data. Printed parental consent forms, which contain identifiable information, will be stored in a locked closet at the PI's office at KU Leuven.*

**How is backup of the data provided?**

*The data will be stored on the KU Leuven central servers with automatic daily back-up procedures.*

**Is there currently sufficient storage & backup capacity during the project? If yes, specify concisely. If no or insufficient storage or backup capacities are available then explain how this will be taken care of.**

*Yes, as the data files are the size of a standard excel document and will never exceed 5 GB. Therefore, we do not anticipate insufficient storage or backup limitations. In case data storage will exceed 5 GB in WP1 due to the images provided by participants, a solution will be worked out in collaboration with KU Leuven ICTS service.*

**What are the expected costs for data storage and back up during the project? How will these costs be covered?**

*Currently, there are no costs expected.*

**Data security: how will you ensure that the data are securely stored and not accessed or modified by unauthorized persons?**

*Both the I-drive and J-drive are fully encrypted and password-protected (only accessible to involved researchers using their employee ID).*

## **6. DATA PRESERVATION AFTER THE FWO PROJECT**

**Which data will be retained for the expected 5 year period after the end of the project? In case only a selection of the data can/will be preserved, clearly state the reasons for this (legal or contractual restrictions, physical preservation issues, ...).**

*All identifiable information will be deleted (i.e., participants' names, email addresses etc.) once the vouchers for compensation and the aggregated results have been sent out, since this information will no longer be relevant or required for study purposes. All other information will be retained for at least 10 years, conform the KU Leuven RDM policy.*

**Where will the data be archived (= stored for the longer term)?**

*The data will be stored on the university's central servers (with automatic back-up procedures) for at least 10 years, conform the KU Leuven RDM policy.*

**What are the expected costs for data preservation during the retention period of 5 years? How will the costs be covered?**

*All datasets will be archived on the servers of KU Leuven. Possible expenses to host these servers are always covered by the research group, but are expected to fall within normal cost ranges (based on the volume of the datafiles).*

## **7. DATA SHARING AND REUSE**

**Are there any factors restricting or preventing the sharing of (some of) the data (e.g. as defined in an agreement with a 3rd party, legal restrictions)?**

*No.*

**Which data will be made available after the end of the project?**

*Anonymized and trimmed datasets of the WPs (.sav format) and data scripts (.R or .sps format) will be available after the end of the project. Participants' personal information (i.e., email addresses) will never be shared.*

**Where/how will the data be made available for reuse?**

*Anonymized and trimmed datasets (i.e., only containing the variables required to reproduce our results; .sav format) and data scripts (.R or .sps format) will be uploaded to the Open Science Framework under a CC-BY license alongside full dataset documentation. Participants' personal information (i.e., email addresses) or audio files will never be shared.*

**When will the data be made available?**

- *Upon publication of the research results.*
- *Previous to the publication for peer-review via a view-only link that is only shared with reviewers.*

**Who will be able to access the data and under what conditions?**

*Study datasets will be made open access via the Open Science Framework and will, before the publication of the study results, be made available to reviewers via a view-only link.*

**What are the expected costs for data sharing? How will the costs be covered?**

*Sharing data on the Open Science Framework is free.*

## **8. RESPONSIBILITIES**

**Who will be responsible for data documentation & metadata?**

*Desiree Schmuck.*

**Who will be responsible for data storage & back up during the project?**

*Desiree Schmuck.*

**Who will be responsible for ensuring data preservation and reuse?**

*Desiree Schmuck.*

**Who bears the end responsibility for updating & implementing this DMP?**

*Desiree Schmuck.*