Who Are the Bad Guys? Disentangling Definitions and Discourses on Terrorism Among Politicians and Citizens

A Data Management Plan created using DMPonline.be

Creator: Amélie Godefroidt

Affiliation: KU Leuven (KUL)

Funder: Fonds voor Wetenschappelijk Onderzoek - Research Foundation Flanders (FWO)

Template: FWO DMP (Flemish Standard DMP)

Principal Investigator: Amélie Godefroidt

Grant number / **URL:** 12D1523N

ID: 198250

Start date: 01-10-2022

End date: 30-09-2025

Project abstract:

Terrorism is intended to change us and the societies in which we live. Yet, despite the important impact of this phenomenon, very little is known about which incidents of political violence are perceived as terrorism by political leaders and the general public. This lack of knowledge is problematic, particularly considering the strategic use and delegitimizing power that comes with attaching the label 'terrorism' to a particular event. In this project, I will develop an innovative theoretical argument to explain how characteristics of the act and actors jointly shape perceptions of and responses to incidents of political violence. I intend to test my argument in two purposefully selected case studies (i.e., Belgium and Norway) and via three steps. First, through conjoint experiments, I will identify the factors that drive politicians and citizens to classify incidents as terrorism. Second, I will systematically analyze how politicians Tweet about political violence and when they decide to use the 'terrorism label' and associated discourses. Third, through a realistic but carefully controlled series of Tweets, I will unravel how such elite frames and counter-frames shape public opinion in the aftermath of violence. By building on insights from terrorism studies and political communication and by using a variety of cutting-edge methodologies, this project will shed an original light on the strategic interplay between actors of political violence, political leaders, and citizens.

Last modified: 25-04-2023

Who Are the Bad Guys? Disentangling Definitions and Discourses on Terrorism Among Politicians and Citizens
Application DMP

Questionnaire

Describe the datatypes (surveys, sequences, manuscripts, objects …) the research will collect and/or generate and /or (re)use. (use up to 700 characters)

The project will generate three primary datasets: (1) one based on two identical conjoint experiments fielded among citizens and politicians in Norway and Belgium, (2) one with Tweets from Norwegian and Belgian politicians from 2011 to 2021, and (3) one based on a 'sequence-of-Tweets' experiment with Norwegian and Belgian citizens. To collect experimental data in Norway, I will apply to the Norwegian Citizen Panel and Panel of Elected Representatives run by the University of Bergen. To collect data in Belgium, I will conduct fieldwork and rely on opt-in samples. Datasets 1 and 3 will be of a rather small volume (i.e., some thousands of respondents), dataset 2 will be of a larger volume.

Specify in which way the following provisions are in place in order to preserve the data during and at least 5 years after the end of the research? Motivate your answer. (use up to 700 characters)

Professor Arnim Langer and I will share responsibility for preserving the data. We will store the data on the KU Leuven network drive of the Centre for Research on Peace and Development. This J-drive is secured and automatically backed up. Following KU Leuven's guidelines on data management, we will store the data for 10 years and will then decide whether it is necessary to store the data any longer. When further storage is no longer deemed necessary, the data will be deleted. Furthermore, upon publication of scientific articles, I plan on sharing a fully anonymized subset of the data for replication purposes on Harvard’s Dataverse.

What’s the reason why you wish to deviate from the principle of preservation of data and of the minimum preservation term of 5 years? (max. 700 characters)

I will not deviate from the principle of preservation of data and the minimum preservation term of 5 years.

Are there issues concerning research data indicated in the ethics questionnaire of this application form? Which specific security measures do those data require? (use up to 700 characters)

The large-N Tweet dataset will contain personal data (in particular the name of the politicians). Although we will remove such data upon publication, we will still be processing personal data and as such we will follow the guidelines that come with that. More specifically, we will comply with both the guidelines of FWO and the Privacy and Ethics Committee of KU Leuven. To do so, we will submit a so-called PRET application (i.e., privacy and ethical clearance) early on in this project (if the project is granted). Such application will also ensure the study conforms to the EU General Data Protection Regulation (GDPR).

Which other issues related to the data management are relevant to mention? (use up to 700 characters)

The experimental dataset will be anonymous as we will not collect any personal data. That is, we will not collect information such as names, (e-mail) addresses, IP addresses, telephone numbers, etc. We will also only collect the most necessary demographic characteristics (such as age and year of birth), so that it will not be possible to identify individuals, either directly or indirectly. Finally, an informed consent form will ensure that the participants in our experiments will have all information about what will happen to their data.

Who Are the Bad Guys? Disentangling Definitions and Discourses on Terrorism Among Politicians and Citizens
DPIA

DPIA

Have you performed a DPIA for the personal data processing activities for this project?

- Not applicable

Who Are the Bad Guys? Disentangling Definitions and Discourses on Terrorism Among Politicians and Citizens
GDPR

GDPR

Have you registered personal data processing activities for this project?

- Not applicable

Who Are the Bad Guys? Disentangling Definitions and Discourses on Terrorism Among Politicians and Citizens

FWO DMP (Flemish Standard DMP)

---

1. Research Data Summary

List and describe all datasets or research materials that you plan to generate/collect or reuse during your research project. For each dataset or data type (observational, experimental etc.), provide a short name & description (sufficient for yourself to know what data it is about), indicate whether the data are newly generated/collected or reused, digital or physical, also indicate the type of the data (the kind of content), its technical format (file extension), and an estimate of the upper limit of the volume of the data.

Dataset Name	Description	New or reused	Digital or Physical	Digital Data Type	Digital Data format	Digital data volume (MB/GB/TB)
S1	Survey experimental data from three different samples: (1) a sample of citizens from the USA (opt-in sample); (2) a sample of citizens from Belgium and Norway (probability sample); and (3) a sample of politicians from Norway (convenience sample). All participants will read a description of a contentious event, in which we manipulate certain characteristics of the event (e.g., whether violence was used). After reading the description, participants will be asked whether they would classify the event as terrorism or not. The experiment will be embedded within an online questionnaire, which will also measure participants' socio-demographics (i.e., observational data).	New data	Digital	Experimental and observational data	.sav, .xlsx, .csv	<100MB
S2	Observational dataset containing Tweets about controversial events tweeted by randomly selected national and local politicians. The variables in the dataset measure specific characteristics of the Tweets (i.e., the labels used in the Tweets), of the events (i.e., whether violence was used), and of the politicians (i.e., their ideology).	New data	Digital	Observational	.sav, .xlsx, .csv	Most likely <1GB
S3	Experimental data from probability samples of citizens from (1) the USA, (2) Norway; and (3) Belgium. In the experiment, we will mimic the communication environment during and after political contestation. The frames and counter-frames used to describe contentious events will be manipulated and participants' responses to such framed information will be recorded. We will also measure participants' socio-demographics (i.e., observational data).	New data	Digital	Experimental and observational data	.sav, .xlsx, .csv	<100MB

If you reuse existing data, please specify the source, preferably by using a persistent identifier (e.g. DOI, Handle, URL etc.) per dataset or data type:

Not applicable.

Are there any ethical issues concerning the creation and/or use of the data (e.g. experiments on humans or animals, dual use)? Describe these issues in the comment section. Please refer to specific datasets or data types when appropriate.

- Yes, human subject data

A Privacy and Ethics application (PRET) has been submitted to the KU Leuven Social and Societal Ethics Committee (SMEC). This PRET application concerns the justifications of the general project, as well as all details of the research design of Study 1 (i.e., the conjoint experiments) and Study 2 (i.e., the Twitter study). Details for Study 3 will be added to this application at a later stage.

Will you process personal data? If so, briefly describe the kind of personal data you will use in the comment section. Please refer to specific datasets or data types when appropriate.

- Yes

The large-N Tweet dataset (S2) will contain personal data, particularly the politicians' Twitter handle (which often includes their names), gender, and political ideology). Although the data will be pseudonymized upon publication, we will still be processing personal data and will follow the guidelines that come with that. To do so, and as indicated above, a PRET application has been submitted (G-2022-5831).

The two experimental datasets (S1 and S3) will be anonymous from the start as we will not collect personal data. We will not collect information such as names, (e-mail) addresses, IP addresses, telephone numbers, etc. We will also only collect the most necessary demographic characteristics (i.e., age, year of birth, socioeconomic status, ethnoreligious background, and political ideology), using broad answer categories. As a result, it will not be possible -not directly nor indirectly- to identify individuals from the data.

Does your work have potential for commercial valorization (e.g. tech transfer, for example spin-offs, commercial exploitation, …)? If so, please comment per dataset or data type where appropriate.

- No

Do existing 3rd party agreements restrict exploitation or dissemination of the data you (re)use (e.g. Material/Data transfer agreements/ research collaboration agreements)? If so, please explain in the comment section to what data they relate and what restrictions are in place.

- No

Are there any other legal issues, such as intellectual property rights and ownership, to be managed related to the data you (re)use? If so, please explain in the comment section to what data they relate and which restrictions will be asserted.

- Yes

In general, publicly posted Tweets can be used as research data, but there are some important considerations to keep in mind.

First, we will ensure that we are complying with Twitter's terms of service and any relevant laws or regulations. Twitter's terms of service state that users own the content they post, but by posting on Twitter, they grant Twitter a license to use, distribute, and reproduce the content. Twitter's terms of service also prohibit the automated scraping of data from their platform without prior permission. Therefore, we have applied to obtain permission from Twitter (i.e., Academic Research access).

Second, when collecting and analyzing Twitter data, it is important to comply with ethical guidelines and ensure that you are not violating the privacy or rights of individual users. As indicated above, we will ensure that we are protecting the privacy and confidentiality of the Twitter users whose content we are using. To do this, we have submitted our plan for review to the ethics committee of KU Leuven. Most importantly, information on the content of the Tweets will be separated from the Twitter handle, which may contain identifiable information. The Twitter handles together with the code to link them back to the data will be stored separately and safely on the secured KU Leuven :I/ drive of the primary investigator (PI). The resulting pseudonymized dataset will be stored in a folder on the KU Leuven OneDrive for Business of the PI. See below for more information about data storage.

2. Documentation and Metadata

Clearly describe what approach will be followed to capture the accompanying information necessary to keep data understandable and usable, for yourself and others, now and in the future (e.g., in terms of documentation levels and types required, procedures used, Electronic Lab Notebooks, README.txt files, Codebook.tsv etc. where this information is recorded).

For all three datasets, the following types of metadata will be generated:

- **a codebook.pdf file**, containing details about
  - the study design, sampling methodology, and response rates;
  - the included variables (i.e., variable labels, exact question wordings, and answer scales with corresponding values);
  - ...
- **a ReadMe.txt file**, containing details about
  - the project and team members (i.e., a short description of the project, the ethical approval number, preregistration doi, and the contact information of all involved researchers);
  - the folders and files in the replication folder;
  - the step-by-step explanation to replicate the results. Note: all r-scripts used to replicate the data will also contain metadata in the form of comments to explain the code;
  - the computational environment used to generate the results. Specifically, using the infoSession() command in R, we will retrieve and report version information about **R**, the operating system, and the attached or loaded packages;
  - the last date of modification and successful replication by the researcher(s);
  - ...
- **a preregistration form**, time-stamped on the website osf.io and using the OSF Preregistration template. This metadata will be available online and get a registration doi, which we will report in the codebook.pdf and ReadMe.txt files.

The **questionnaires** will also be available for the survey-based datasets (S1 and S3).
Finally, this **Data Management Plan** as well as the **ethical clearance** (PRET application) will also be saved as metadata in the folder of this project.

Will a metadata standard be used to make it easier to find and reuse the data? If so, please specify (where appropriate per dataset or data type) which metadata standard will be used. If not, please specify (where appropriate per dataset or data type) which metadata will be created to make the data easier to find and reuse.

- Yes

- **codebook.pdf:** We will follow the KU Leuven guidelines and, hence, use DDI to create the codebook (https://ddialliance.org/training/getting-started-new-content/create-a-codebook)
- **ReadMe.txt:** We will use the KU Leuven template (https://www.kuleuven.be/rdm/en/guidance/documentation-metadata/README)
- **preregistration**: We will use the OSF Preregistration template

3. Data storage & back-up during the research project

Where will the data be stored?

The **experimental dataset (S1 and S3)** will be stored in a folder on the **KU Leuven OneDrive for Business of the primary investigator (PI)** . OneDrive for Business is suitable for strictly confidential data, as long as multifactor authentication with the KU Leuven Authenticator app is activated. OneDrive is also suitable to save human subject data in this case, since the data will be anonymized. Access to these datasets will be restricted to the PI and only the PI can decide to grant access to collaborators.
The **Twitter dataset (S2)** will also be stored in a folder on the **KU Leuven OneDrive for Business of the PI**. As this dataset will contain personal data (i.e., Twitter handles, which often include names), extra precautions will be taken:

- The dataset will be *pseudonymized.* We will replace all identifiable information (e.g., Twitter handles) with a pseudonym and code indirectly identifiable information (e.g., age) in a way that makes it impossible to re-identify individuals. A separate record, or so-called coding table, will be created that links the pseudonyms to the identifiable information. This coding table will be kept securely in the **personal KU Leuven I-Drive of the PI**, and thus separated from the pseudonymized data (which will be stored on OneDrive), to prevent re-identification.
- We will *strictly manage access rights* to this dataset. Specifically, we will check regularly whether only the right people have access to the data (need-to-have) and whether these people only have access to the data they strictly need (need-to-know).

In case any of the datasets will be used in a collaborative project among KU Leuven staff, this dataset will also be stored in a folder on a **shared KU Leuven J-Drive**. The shared network drive is based on the standard file storage solution of KU Leuven and is suitable for (strictly) confidential data as well as for storing and processing all types of personal data.

How will the data be backed up?

Files in OneDrive are automatically backed up, up to 100 versions per file. A second copy will be kept on the KU Leuven I:\ network drive (i.e., the

personal network drive) of the primary investigator.

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

We consider that there will be sufficient storage and backup capacity during the project. OneDrive for Business has a standard storage capacity of 2 TB but this can be extended to 5 TB upon motivated request through ICTS Service Point. The estimated size of all datasets combined is < 1.5 GB.

How will you ensure that the data are securely stored and not accessed or modified by unauthorized persons?

- We will **strictly manage access rights** to datasets. Only the PI will have access to the datasets and will have to grant access to anyone else.
- We will activate **multifactor authentication** with the KU Leuven Authenticator app for OneDrive Business.
- The laptop(s) on which the datasets are stored are **password-protected**.
- Personal data will be **pseudonymized** and the coding table, which makes it possible to link the identifiable information with the pseudonymized data, will be stored on an encrypted KU Leuven I-drive.

What are the expected costs for data storage and backup during the research project? How will these costs be covered?

There are no expected costs for data storage, as OneDrive for Business is free for staff and students of KU Leuven. The I- and J-drives are also provided (and secured) by KU Leuven.

4. Data preservation after the end of the research project

Which data will be retained for at least five years (or longer, in agreement with other retention policies that are applicable) after the end of the project? In case some data cannot be preserved, clearly state the reasons for this (e.g. legal or contractual restrictions, storage/budget issues, institutional policies...).

The three datasets will be preserved for minimally 10 years after study completion, in line with the KU Leuven RDM policy.

Where will these data be archived (stored and curated for the long-term)?

Upon completion of this project, all datasets, along with the documentation and metadata necessary to reuse the data, will be transferred to the KU Leuven secured KU Leuven K:\ network drive managed by the research group's leader (Prof. dr. Arnim Langer).
Upon completion of a publication, the dataset used in that publication, along with the documentation and metadata necessary to reuse the data, will be deposited at the Harvard Dataverse and the **KU Leuven RDR**. Both are generic, online data repositories where one can publicly share and preserve research data. While the Harvard Dataverse is well-known in my field of study (and often used by renowned journals), the KU Leuven RDR is the preferred option for KU Leuven research data publication as it's the university's institutional research data repository.

What are the expected costs for data preservation during the expected retention period? How will these costs be covered?

There are no expected costs for data preservation.

5. Data sharing and reuse

Will the data (or part of the data) be made available for reuse after/during the project? In the comment section please explain per dataset or data type which data will be made available.

- Yes, in an Open Access repository

The three collected datasets (in their entirety) will be made available upon publication of the research results.

If access is restricted, please specify who will be able to access the data and under what conditions.

Not applicable.

Are there any factors that restrict or prevent the sharing of (some of) the data (e.g. as defined in an agreement with a 3rd party, legal restrictions)? Please explain in the comment section per dataset or data type where appropriate.

- No

Where will the data be made available? If already known, please provide a repository per dataset or data type.

Upon completion of a publication, the dataset used in that publication, along with the documentation and metadata necessary to reuse the data, will be deposited at the Harvard Dataverse and the **KU Leuven RDR**. Both are generic, online data repositories where one can publicly share and preserve research data. While the Harvard Dataverse is well-known in my field of study (and often used by renowned journals), the KU Leuven RDR is the preferred option for KU Leuven research data publication as it's the university's institutional research data repository.

When will the data be made available?

Upon publication of the research results.

Which data usage licenses are you going to provide? If none, please explain why.

CC-BY-NC-SA-4.0:

- Free to share and adapt.
- Give appropriate credit, and indicate if changes were made.
- Do not use the material for commercial purposes.
- Distribute your contributions under the same license as the original.

Do you intend to add a PID/DOI/accession number to your dataset(s)? If already available, you have the option to provide it in the comment section.

- Yes

The Harvard and KU Leuven repositories use a DOI.

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

There are no expected costs for data sharing, as sharing datasets in the selected data repositories is covered by the repositories.

6. Responsibilities

Who will manage data documentation and metadata during the research project?

Amélie Godefroidt (PI)

Who will manage data storage and backup during the research project?

Amélie Godefroidt (PI)

Who will manage data preservation and sharing?

Amélie Godefroidt (PI)

Who will update and implement this DMP?

Amélie Godefroidt (PI)