Data management plan

Anna Durnová, Sarah Helena Schäfer, & Zita Zeberer

CiDAPE

Climate, Inequality & Democratic Action:
The Force of Political Emotions

Funded by
the European Union

Disclaimer

This project has received funding from the European Commission under the HORIZON Research and Innovation Actions Grant, No 101132327. The document reflects only the authors' views. The European Union is not liable for any use that may be made of the information contained herein.

Dissemination level

PU - Public

Type

DMP – Data Management Plan

CIDAPE
Climate, Inequality, and Democratic Action: The Force of Political Emotions

Data management plan

Deliverable D1.2

Authors: Anna Durnová[1], Sarah Helena Schäfer[1], & Zita Zeberer[1]

[1] University of Vienna

Due date: August 2024

Executive Summary

The guiding principles of CIDAPE are data protection, good scientific practice, and open access policy. All institutions involved in CIDAPE comply with European and national legislation and fundamental ethics principles, including those reflected in the Charter of Fundamental Rights of the European Union and the European Convention on Human Rights, its Supplementary Protocols and General Data Protection Regulation (Regulation (EU) 2016/679).

Personal data will never be published other than in an aggregated and/or anonymous form to protect the privacy of those concerned (in interviews, groups discussions and panels / both offline and online). Generally, all personal usage data (from either individual or collective interactions) will be dealt with based on three main principles – the volunteering principle, the consent form, and privacy considerations. With interviews of citizens involved in particular activities, where the concrete institutional embedding is necessary for the context of the concrete case, there will be negotiations with the data suppliers to specify in what form the data can be used, either as background information only or as directly quotable information.

If at any point respondents want to withdraw their response, they will be welcome to do so. All data sets will be stored securely on a password encrypted device, and when provided by the research infrastructures, secured cloud storages will be used.

1 Data Management and Data Protection

CIDAPE promotes the responsible management of research data and outputs, that are outlined in the Data Management Plan (DMP). This living document aims to outline from the start of the project all aspects of the lifecycle of the research data (collection, access, storage, sharing, deletion etc., both during and after the project). To keep this document up to date, the latest developments will be recorded continuously, as the project and the several different data collection efforts progress.

1.1 Data protection

CIDAPE collects data from social media, activists, stakeholders, and citizens throughout the lifetime of the project. Confidentiality of respondents will be scrupulously protected. CIDAPE ensures that data will and can only be used for the exact purpose of the respective study. Personal data will never be published other than in an aggregated and/or anonymous form to protect the privacy of those concerned (in interviews, focus groups, and group discussions). With expert interviews, where the concrete institutional embedding is necessary for the context of the concrete research, there will be negotiations with the data suppliers to specify in what form permission is granted for use, either as background information only or as directly quotable information.

1.2 Good scientific practice

CIDAPE applies accepted rules of good scientific practice throughout the project. CIDAPE will appropriately attribute the sources of cited material and not infringe on the intellectual property rights of third parties. All legal aspects pertaining to the management of Intellectual Property (IP) will be comprehensively governed by the Consortium Agreement. Furthermore, strict ethical standards for research, as used in the international community with respect to the use of sensitive data will be adhered to. As for the publications of scientific articles, all publications that stem directly from CIDAPE's research will be attributed to CIDAPE and the affiliation of the consortium partners will be used. All CIDAPE project members will be named and identified with their affiliations in dissemination outputs of CIDAPE such as videos, dramatising content, public debates, stakeholder forums and the like.

1.3 Research Data Management

CIDAPE adheres to the FAIR Principles, which includes ensuring that the data is findable by sharing it through the electronic infrastructure of the project and trustworthy repositories, such as GESIS, Zenodo, or Mendeley Data. These repositories will assign Digital Object Identifiers (DOIs) to ensure that the dataset is identifiable and citable. The data will be gathered in accordance with all legal and ethical standards and will only be made

accessible for further research purposes under a Creative Commons license unless a public license is already implemented by the repository. If the data is used as a basis for a publication, access will be granted as soon as the publication becomes public. Data that is not linked to publications will be made available after a short internal grace period of three months after its generation. Prior to granting access, partners will be asked if they prefer not to give immediate access for reasons such as planning a publication based on the data, or if the data is pertinent to vital interests. The data can be accessed using standard software in the relevant disciplines, and metadata will be linked via the repositories' interface. The interoperability of the data will be ensured by adding descriptive metadata and keywords in accordance with the ontology of the relevant discipline, and by organising folders and files in a clear and identifiable structure using a common naming convention. CIDAPE strives to make the data reusable by publishing articles in peer-reviewed journals and by pre-registering the studies on the project website to ensure that the data complies with scientific quality standards. This includes allowing for replication, cross-comparison, and re-testing against known misconceptions.

2 Data Management in the different Work Packages

In this section the data management practices of all Work Packages are detailed, followed by a table for a complete overview about the management of the research data in CIDAPE.

2.1 Work Package 1

2.1.1 Data collection and summary

There is no primary data collection in Work Package 1 and sensitive personal data is not collected. The main theoretical output of this WP is the creation of a "Theory guide on structural embeddedness of emotions" which is a literature review of publicly available secondary literature on climate change and emotions. Additionally, the creation of a "Methodological guide to work with emotions" using innovative approaches and ideas from the CIDAPE consortium as well as relevant secondary literature on how to analyse emotions effectively in the social sciences. Contact data of consortium partners is collected for administrative purposes in line with the Consortium Agreement.

2.1.2 Open Science practices

WP1 pursues an open-access policy for both scientific and practical outputs. Public deliverables of the theory- and methodological guides will be publicly accessible on the project website, among others.

2.1.3 Data storage and data security

Data resulting from the administration and management of the consortium is stored on the u:cloud storage, a secure cloud storage provided by the University of Vienna and shared among consortium members. Documents, deliverable reports and information are shared with the consortium primarily via E-mail and Slack (version for educational institutions).

2.2 Work Package 2

2.2.1 Data collection

- WP2 starts data collection for the Policy Document Analysis in 2025. Data management has been outlined already and is detailed below.
- Joint Data collection jointly by WP2 and WP3 on social media communication (Facebook, Instagram, TikTok) during EU elections on climate policy communication and emotions.

2.2.2 Data summary

- Social Media media communication by public actors (joint data WP2 and WP3)

Publicly available user information and data (posts, videos, comments) posted across Europe, focusing especially on the European level and zooming in on specific country cases and languages, content related to climate change related policies (e.g., European Green Deal) and climate change related political events (e.g., elections,

protests). This includes Social Media communication from political parties, their candidates, MEPs of the European parliament and of climate and environmental organisations and media outlets. Tik Tok developer API; Facebook and Instagram via Crowdtangle.

- Policy Document Analysis

WP2 will only collect publicly available data such as parliamentary minutes, pertinent policy documents, and newspaper coverage surrounding select climate policies. Only data that do not contain personal information and do not require anonymisation will be collected.

2.2.3 Open Science practices

- Social Media media communication by public actors (joint data WP2 and WP3)

Open Code: Code will be made available on Github (or another repository). Adhering to open science principles, metadata of generated datasets, including user IDs, video IDs, and comment IDs studied, may be publicly deposited, enabling other researchers to replicate our work after the end of the project. CIDAPE research will here follow EU guidelines (i.e., GDPR No 2016/679) as well as the university's ethical guidelines, and share only what is in line also with the TikTok Developer Terms, and the Terms by Crowdtangle. The research project has been approved by the ethics committee of the University of Vienna.

2.2.4 Data storage and data security

- Social Media media communication by public actors (joint data WP2 and WP3)

Data is stored in csv and json format on the secure cloud storage provided by the University of Vienna. Personal data such as usernames, locations, and posted contents is in general first pseudonymised. Confidentiality of all users will be scrupulously protected. CIDAPE ensures that these data will and can only be used for the exact purpose of the project.

- Policy Document Analysis

Data will be stored securely on a cloud storage provided by Bielefeld University.

2.3 Work Package 3

2.3.1 Data summary

- Dataset 1: published scientific articles in csvs format with meta data about the scientific articles, the scientific articles are generated and stored in PDF format.
- Dataset 2: approximately 10,000 comments and 100 posts per language (German, English, Spanish) posted on TikTok in csvs format with posts, comments and related metadata, video files of manually selected videos.
- Dataset 3: social media communication on climate change by citizens across different social media platforms TikTok: in csvs format with posts, comments and related metadata as collected via the TikTok Developer API; video files of selected videos Facebook/Instagram: data formats not clear yet. They depend on the access possibilities of the new Meta content library.
- Dataset 4: social media communication by public actors (see details “joint data WP2 and WP3” above).

2.3.2 Open Science practices

- Open Code: Code will be made available on Github (or another repository).
- Adhering to open science principles, metadata of generated datasets, including user IDs, video IDs, and comment IDs studied, may be publicly deposited, enabling other researchers to replicate our work after the end of the project. CIDAPE research will here follow EU guidelines (i.e., GDPR No 2016/679) as well as the university's ethical guidelines and share only share data that is in line with the Developer Terms by the Social Media Platforms.

2.3.3 Data storage and data security

- Using the secure cloud storage provided by the University of Vienna.

Personal data such as usernames, locations, and posted contents is in general first pseudonymised. Confidentiality of all users will be scrupulously protected. CIDAPE ensures that these data will and can only be used for the exact purpose of the project.

2.4 Work Package 4

2.4.1 Data collection

In WP4, 30 semi-structured interviews in each of the four countries (Norway, Spain, Slovakia, and Austria) will be conducted, so 120 interviews in total. There will be 3 focus group discussions conducted in each country, so 12 in total. Interview data collection will start in September 2024. Group discussions will be held in early 2025.

2.4.2 Data summary

Interviews and group discussions will be recorded and transcribed. The resulting data from interviews and focus group discussions will take various forms, including audio recordings, transcripts, and demographic information pertaining to all participants and to interviewers.

To ensure the privacy and anonymity of all participants, the interview and group discussion data will be anonymised. Every participant is assigned a unique code, ensuring anonymity as it prevents the direct association of personal data with individual identities. This code is stored independently from the pseudonymised data to maintain confidentiality and privacy.

2.4.3 Open Science Practices

To ensure the privacy and anonymity of all participants, the data will not be made publicly available. The analysis of the interviews and group discussions, translated to English, will be made available as part of research articles that will be published Open Access.

2.4.4 Data storage and data security

Stringent measures will be taken to protect the stored data; it will be securely stored in an encrypted cloud service throughout the project's duration. Backups of the data will be held in encrypted offline data storage, protecting it from unauthorised access by third parties. Data protection measures will be implemented, ensuring the safe storage of this information with no access granted to third parties.

Audio recordings, anonymised/encrypted transcripts and the lists of demographic variables of all participants and interviewers will be stored in secured cloud storage provided by the University of Vienna. Personal data gathered will be retained for the project's duration and as long as it supports the creation of scientific outputs like publications, conference papers, and qualification papers, but not beyond. Once the data has been aggregated and individual access to personal data is deemed unnecessary, it will be deleted. Transcripts and audio files from interviews and group discussions will be digitally archived in folders that are exclusively accessible to project team members.

2.5 Work Package 5

2.5.1 Data collection

The aim of WP5 is to understand farmers' perceptions of climate change norms. Ethnographic observation and narrative go-along interviews will be conducted (n= approx. 10-15 to reach saturation of topics; duration of one interview approx. 1,5 hours), in some cases supplemented by semi-structured interviews. The data will be collected in two selected regions of South Moravia in Czechia and Veneto in Italy.

2.5.2 Data summary

Interview transcripts in text files (Word/PDF). The interviews will be carried out with farmers in Czechia and Italy. The transcripts will include basic identification data: nicknames of informants, the place, date, time and length of the interview, gender and age of informants and interviewers. The interviews will be stored in original

languages (Czech and Italian), however, translated summaries of the analysis with examples/excerpts will be provided in English.

2.5.3 Open Science Practices

The go-along interviews will focus on the opinions, attitudes and practices of residents of small municipalities in two particular Czech and Italian regions. Due to the narrow delineation of the research field and the focus of the interviews on relationships within the community (differing perceptions of climate change), publishing the full transcripts of the interviews cannot maintain the anonymity of the informants. Therefore, for ethical reasons, the transcripts will not be stored publicly outside of CIDAPE members. The analysis of the interviews, translated to English, will be a part of a research article that will be published Open Access.

2.5.4 Data storage and data security

Audio recordings, anonymised/encrypted transcripts and the lists of demographic variables of all participants and interviewers will be stored in secured cloud storage provided by the Metropolitan University Prague. The informed consents will be stored in offline data storage at MUP. The interview transcripts will be analysed using the qualitative data analysis programme Requal, produced by Charles University (https://requal.fsv.cuni.cz). A server-based version of the programme will be used to enable online teamwork. The coded data will be stored on the servers of Charles University.

2.6 Work Package 6

2.6.1 Data collection

In the framework of WP6, there will be 4 citizen expert panels organized (2 in Czechia and 2 in Italy) during 2025. Preparation for the data collection efforts is ongoing.

2.6.2 Data summary

- Demographic variables of all participants.
- Demographic variables of all presenters and facilitators.
- Transcription and recording of presentations and Q&A sessions.
- Transcription and recording of CEPs participants' discussions with facilitators.

2.6.3 Open Science practices

Open science practices play a key role and will be achieved in collaboration with the Czech Social Science Data Archive and the Italian Open Science Support Group. This plan will then be submitted to an Ethical Research Committee of Charles University and the joint ethical committee for research of the SNS.

2.6.4 Data storage and data security

Stringent communication toward the Czech Social Science Data Archive to follow best practices when designing the data collection is foreseen.

2.7 Work Package 7

Work Package 7 in collaboration with Utrecht University and People in Need (PiN) adheres to the Data Protection Policy in use by their organisations, which provide an operational framework on how to work with (personal) data.

2.7.1 Data collection

- Two roundtables of alternative citizen engagement (PiN in CZ and SK).
- Interviews for case studies, participant observations and diary entries will be collected by Utrecht University.

2.7.2 Data summary

- Roundtable participants' personal data (name, contact info such as email or telephone number, their organizational background) will be collected via an attendance sheet, which will come along with an informed consent detailing for what purpose the personal data is collected, how will it be stored and used and how their consent can be withdrawn and the processing of the data stopped. Transcriptions and recordings of the roundtable discussions will be made with the informed consent of all participants.
- Three interviews case study, three participant observation and diary entries (in text format .sav).

2.7.3 Open Science practices

- The data collected during the roundtables is not to be openly accessible, since it is marked as "sensitive" by the grant agreement – it can only be shared among the participants of the activity if it is required to fulfil the goals of the project – and also because it would be a breach of the requirement to not publish personal data in any form other than an aggregated and/or anonymous one in order to protect the privacy of those concerned. Only the anonymised version of the discussions' transcriptions may be used for the purpose of implementing the project.
- Data collected during the case study interviews, participant observations and diaries will be made findable, accessible, interpretable and reusable. The data will be published via YoDa. Upon publication, YoDa assigns a Digital Object Identifier (DOI) to the data package, making it Findable and citable for the research community at large. Given the data is personal and sensitive, it will be published under restricted access. This means while the metadata and specific folders/files to do with documentation will be openly published, the data files themselves will be made available only upon request.

2.7.4 Data storage and data security

- Personal data collected by PiN is stored only for the necessary period of time. This is based mainly on the Act on Archiving and the Act on Accounting and is described more in detail in the internal regulation of the organisation. Once a year data is reviewed, and such data whose storage period has passed are removed, unless a legitimate ground for their continued storage exists. The data is stored on secured cloud servers and/or in our archiving system ELO. In both cases the access to the data is restricted according to relevance to specific persons (for example: an employee who has nothing to do with the CIDAPE project will not be able to access the data freely).
- During the study at Utrecht University, both SURFdrive and the O: drive at FSBS are used for data storage. SURFdrive is utilized for day-to-day data management and data preparation. The processed data would then be placed in the O: drive. SURFdrive is a cloud storage solution provided by SURF for Dutch education and research institutions. SURFdrive allows users to share and collaborate on files within a secure community cloud. The information security within SURFdrive meets the standards of Dutch and European privacy legislation. The data is always sent encrypted over networks. SURFdrive has a 30- day backup & recovery facility. The O: drive is a networked drive within Utrecht University, available to all employees and students. Users only have access to their faculty's O: drive and specific project folders there. As the study is completed, data will additionally be stored in YoDa. Any personal data will abide GDPR standards. This includes secured storage, access control and pseudonymization
- Utrecht University is the controller of the above-mentioned data. Within the organization, the PhD of the research project is primarily responsible for overseeing the management and processing of the data. All datasets will be preserved for 10 years, according to the policy framework research data of the Utrecht University. Legal basis by which data is collected is based on informed consent.

2.8 Work Package 8

2.8.1 Data collection and summary

For social media accounts Instagram and Twitter as well as for the CIDAPE website, WP8 will collect milestones from all WPs and centralise public dissemination for CIDAPE. Both a "Working paper on Ethical issues

involving positionality, bias", "Guidelines to use participatory strategies" and a "Publication and dissemination plan" created by WP8 will serve as guidelines for the consortium. For ethical guidelines and participatory strategies, we will use secondary literature. For the publication plan we will determine relevant academic journals in which we want to publish CIDAPE research articles. At the end of the project, WP8 provides a "Report on Democratic Action and the Force of Political Emotion" using CIDAPE research results from WP1-WP7 over the course of our project.

2.8.2 Open Science practices

Open science practices, ethical standards and intellectual property rights related to investigations and data management in WP2-WP7 are overseen by this work package. WP8 adheres to established ethical standards outlined in European and national laws and follow best practices for research collaboration and data management and oversees and coordinates the application of ethical standards of all disciplines involved in the project.

2.8.3 Data storage and data security

News on CIDAPE progress and milestones will be made publicly available on Instagram, Twitter and the CIDAPE website. Working papers and guidelines will be stored on the 'u:cloud' cloud storage service provided by the University of Vienna and shared among consortium members primarily via E-mail or Slack, just as well as working papers. The "Report on Democratic Action and the Force of Political Emotion" will be published in a relevant academic journal (e.g., Policy Sciences). WP8 will further oversee strategies for outreach work in form of policy briefings or media coverage and engage with target audiences in the civil, stakeholders, and policy and public administration sectors through communicating CIDAPE milestones in form of press releases and interviews, and project newsletters. Together with the Executive Board and the Coordinator, WP8 prepares the content and timing of "Communication, Dissemination, Open Science and Visibility" activities, in line with the Consortium Agreement.

TABLE 1: MANAGEMENT OF THE RESEARCH DATA

Method	Type of Data, Data Format	Data Sharing, Open Access	Data Storing, Data Protection
Interviews (WP4, WP5, WP6)	Audio Recordings, Transcripts, Demographic variables of all participants, Demographic variables of interviewers	Interview data will be anonymised	Safe storage with no access to third parties. Data will be stored in an encrypted cloud service for the duration of the project. The backup of the data will be stored at encrypted offline data storage.
Group Discussions (WP4)	Audio Recordings, Transcripts, Demographic variables of all participants	Group discussion data will be anonymised	Safe storage with no access to third parties
Ethnography (WP3, WP5, WP6)	Participatory research: interviews, recordings and videos, digital ethnography in WP3	Data repositories Dryad, Figshare, re3data, Social Impact Open Repository (SIOR), or OpenAIRE's Zenodo	
Policy Document Analysis (longitudinal) (WP2)	Policy documents, parliamentary minutes, legal texts, media articles, social media posts	Publicly available data	Only data that do not contain personal information and do not require anonymisation will be collected for these tasks of WP2

Text-as-data / text discovery (e.g., topic modelling) (WP3)	Data Tables (R, Python). Using the APIs of social media platforms, WP3 will collect user account meta data and textual data. Most relevant meta data for this WP are ID, name, username, profile bio, location. On the level of texts (e.g., tweets) most important meta data are mentions, retweets, and the content. The study uses only public social media data.	In general, WP3 will only display aggregated information (like a heat map) in publications. For Twitter, in accordance with the Twitter developer policy, WP3 may share the collected Twitter IDs (no contents) with other researchers to facilitate peer-review and validation. Open Code: Code will be made available on Github (or another repository)	Data is stored in csv and json format on the secure cloud storage provided by the University of Vienna
Deliberative Methods (WP2, WP6, WP7)	Demographic variables of all participants, demographic variables of all presenters and facilitators, transcription and recording of presentations and Q&A sessions provided by civil society members, transcription and recording of participants' discussions with facilitators, final document written by members of research involving group discussions		Data will be stored in an encrypted cloud service for the duration of the project. The backup of the data will be stored at encrypted offline data storage.