

Report on the outcomes of a Short-Term Scientific Mission¹

Action number: CA21129

Grantee name: Mariken A.C.G. van der Velden

Details of the STSM

Title: Populist Opinions for the Young

Start and end date: 15/01/2025 to 19/01/2025

Description of the work carried out during the STSM

During the STSM, significant progress was achieved in multiple areas aligned with the initial working plan. The primary focus was on data collection, annotation, and preliminary model testing, along with foundational theoretical groundwork. The activities carried out are summarized as follows:

1. **Data Collection Strategy and Initial Dataset Compilation:** A comprehensive strategy was developed to collect data from TikTok and YouTube, focusing on content from political parties in the Netherlands (NL) and Germany (DE) during the European Parliamentary Elections 2024 (EPE24). The first batch of data was successfully gathered, ensuring a representative sample for subsequent analysis. This step was essential to establish a solid empirical basis for the project.
2. **Gold Standard Annotation for Reactionary Populism Concepts:** A gold standard dataset was created by annotating textual content for seven distinct concepts that collectively define reactionary populism. This annotation process was rigorous and based on predefined guidelines, ensuring high reliability and consistency. The annotated dataset serves as a critical resource for training and evaluating machine learning models.
3. **Testing a Generative AI Model:** The Generative AI (GenAI) model was tested for its ability to detect and annotate reactionary populism within text. Preliminary results demonstrated the model's potential, although additional fine-tuning and evaluation are necessary for optimizing its performance. These tests provided insights into the model's strengths and limitations, forming the basis for future improvements.
4. **Development of Theoretical Framework and Introduction Structure:** The theoretical framework was established to contextualize the study of reactionary populism within contemporary political discourse. This framework integrates existing literature with novel insights derived from the data analysis. Additionally, the structure for the introduction of the final report or publication was set up, outlining key objectives, hypotheses, and methodological approaches.

¹ This report is submitted by the grantee to the Action MC for approval and for claiming payment of the awarded grant. The Grant Awarding Coordinator coordinates the evaluation of this report on behalf of the Action MC and instructs the GH for payment of the Grant.

While the majority of activities proceeded as planned, minor adjustments were made to optimize workflow efficiency. For instance, the scope of the data collection was initially broader but was refined to focus exclusively on TikTok and YouTube due to their relevance and data availability. These changes ensured the project's objectives were met without compromising quality.

Description of the STSM main achievements and planned follow-up activities

The collaborative effort with the RPTU team ensured progress in data collection, annotation, and pipeline optimization. Key achievements and their alignment with the planned outcomes are detailed below:

1. **Pipeline Optimization and Data Collection:** RPTU team resolves issues in the snowballing pipelines, streamlining the process for generating robust datasets. By February 4, proper material for Germany (GER) and the Netherlands (NL) was collected, meeting the planned timeline and ensuring the availability of high-quality data.
2. **Annotation and Variable Selection:** RPTU team selected 50 examples classified by GPT for relevant variables by January 22. These were sent to Mariken, who uploaded them onto Annotinder for streamlined annotation. By February 4, Michaela, Mariken and Talisa completed the annotation of these 50 examples, adhering to the timeline and ensuring the reliability of the gold standard dataset.
3. **Literature Review Supervision:** RPTU team supervises the research assistants responsible for collecting literature, ensuring comprehensive coverage of relevant theoretical and empirical work. This contributed to the depth and rigor of the project's conceptual framework. We have reserved February 25 and 26 for intensive, full-day collaborative work sessions.
4. **Collaborative Work and Future Planning:** On February 5, the team convened on Zoom to review progress and outline subsequent tasks. Michaela facilitated the meeting by providing the Zoom link. Following the meeting, RPTU team provides Talisa, Mariken and Michaela with an additional 100 texts from the new snowball sample (the final GPT test) for immediate annotation. This ensured that the project continued to progress efficiently.
5. **Data Expansion and Intensive Work Sessions:** RPTU team generates llama-data by February 24, contributing additional insights and expanding the dataset. The team reserved February 25 and 26 for intensive, full-day collaborative work sessions. The work commenced with a Zoom meeting on February 25 at 10 am to synchronize efforts and finalize tasks.

The team agreed to maintain close collaboration, building on the momentum established during the STSM. Future plans include:

- Extending the annotated dataset to include additional variables and concepts.
- Conducting further tests with Generative AI models to refine classification accuracy.
- Preparing publications to disseminate findings and methodological insights.