

## **Data collection notice for DE-CONSPIRATOR**

This data is collected in the public interest as part of the research project DE-CONSPIRATOR, a project funded by the European Union Horizon programme (Grant No. 101132671), conducted by Dr Bharath Ganesh at the University of Amsterdam, Institute for Logic, Language, and Computation.

### **Purpose of the research project**

The research will be focused on detecting, identifying, and studying networks that spread Russian and Chinese “foreign information manipulation or interference” (FIMI), such as misinformation, fake news, and disinformation, or engage in information suppression within the EU and partner countries. The research will collect data on social media platforms and websites. Specifically, the consortium will provide the researcher with identifiable accounts associated with FIMI, for example accounts run by Russian state outlets, pro-Kremlin or pro-China influencers engaging in information manipulation, and other actors.

Using APIs and web scraping techniques, the research project will collect comprehensive data on these users as well as any hyperlinks, shares, reposts, and other digital trace data. This will also include rich multimedia data (eg. videos, images, etc). These data will be used to construct networks of FIMI influence within the EU and partner countries.

To collect data, we will use APIs including the Telegram API, YouTube API, the X API, Meta platforms, and the TikTok Research API (which we access through legally mandated means by the EU Digital Services Act application or from open APIs). Using these APIs as well as open-source web scraping techniques, we aim to reconstruct the information environment surrounding Russian and Chinese FIMI and the transnational and European networks that these actors rely on.

In addition to user-based queries, we will search hashtags and keywords, such as those related to elections, in order to identify users and websites to feed into the data collection procedures outlined above, supplementing the accounts provided by the rest of the DE-CONSPIRATOR consortium.

### **What data is collected?**

- Public posts on Telegram, X, TikTok, Facebook, Instagram, or other social media platforms that are posted by accounts engaged in FIMI (based on expert validation) or accounts engaging on keywords or hashtags related to elections and key issues (eg. global conflicts, migration, protests, etc).
- Data on users that have liked or shared posts on social media platforms in order to identify potential automated or spam activity.
- Data collected from public websites, alternative media, and blogs.
- Data on accounts and websites shared on social media platforms from accounts that we collect, such as hyperlinks/URLs and shared channels, users, and organizations.
- Advertisements placed on social media platforms

We expect the data to number in the millions of posts and thousands of users. As this is sensitive data, we will use the highest standards of data protection to maintain the privacy of users whose data is collected, outside of justifiable identification of users (eg. public figures, influencers with thousands of followers, organizations rather than persons).

### **Discomfort, Risks & Insurance**

There are no risks to participation in this research beyond existing risks that may result from publicly posting on social media platforms, where data is routinely collected and investigated. Our standards of data protection mitigate potential risks.

For all research at the University of Amsterdam, a standard liability insurance applies.

### **Confidential treatment of personal data**

The information gathered over the course of this research will be used for the purpose of this research project. Personal details will not be used in publications outside of justifiable identification of users where needed to answer the research questions.

Data will be pseudonymised for processing (outside of justifiable users that are publicly visible, eg. public figures, organizations, and users with thousands of followers). The pseudonymisation key will be encrypted with a password only known to the lead researcher and will only be accessed by the lead researcher.

Data will be retained for a minimum period of 10 years in a private datastore. Aggregate data, which will be stripped of personal information will be made publicly available through publication. Personal data will only be stored as long as is necessary for the research and will be deleted as soon as possible.

### **Further information**

For further information on the research project, please contact Dr Bharath Ganesh, phone number: +31 20 525 2980, or email [b.ganesh@uva.nl](mailto:b.ganesh@uva.nl); or by contacting Spuistraat 134, 1012VB Amsterdam.

If you have any complaints regarding this research project, you can contact the secretary of the Ethics Committee of the Faculty of Humanities of the University of Amsterdam, [commissie-ethiek-fgw@uva.nl](mailto:commissie-ethiek-fgw@uva.nl); Binnengasthuisstraat 9, 1012 ZA Amsterdam, The Netherlands.

For more information on your rights as a data subject if your personal data is used in this research, you may visit the following website:  
<https://extranet.uva.nl/en/content/a-z/privacy/what-are-my-rights/what-are-my-rights.html>