

LAGO Data Protection Notice

Information for the processing of personal data in accordance with art. 14 GDPR

The purpose of this data protection notice is to inform data subjects about the processing of their personal data. Considering the technical nature of the module and limitations imposed by the research design (i.e., scale), it is considered that informing those data subjects directly would involve a disproportionate effort. For this reason, this information is made publicly available via the project's website in accordance with art. 14 GDPR and with its potentially applicable derogations (art. 14 (5) (b) GDPR¹), as an effort of enabling the data subjects to be informed about the data processing and to exercise their rights. This notice refers to the specific module of the LAGO responsible for collection of data from online sources.

Data will be collected from social media posts (publicly available) in particular from X (Twitter), relevant to the Fight Crime & Terrorism (FCT) domain.

1. The Project

[LAGO](#) will deliver the foundation for a trusted EU FTC Research Data Ecosystem (RDE) to address the so-called "Data Issue" in the FCT research landscape, i.e., the lack of domain-specific data in sufficient quality and quantity to enable appropriate training and testing of the developed methods, platforms and tools. LAGO will be instrumental in identifying common barriers and subsequently providing the structural, governance and technical foundations to foster and innovate data-oriented research collaboration among LEAs, security practitioners, relevant EU agencies, academic and industry researchers, policy makers and regulators. For this purpose, LAGO will develop an evidence-based and validated multi-actor Reference Architecture for the FCT RDE for these actors to deposit, share and co-create data and tools for FCT research purposes based on common rules, protocols, standards and instruments in a trusted and secured environment. The envisaged Reference Architecture and accompanying governance framework will be based on the design principles of decentralisation, data sovereignty, data quality, openness, transparency and trust and comply with EU values and principles on data protection, privacy and ethics. The Reference Architecture will be accompanied by a TRL-7 Reference Implementation of added-value technological tools to ensure practical realisation of the Reference Architecture as multiple data spaces and across the full range of concrete usage scenarios. A Roadmap will finally provide the consolidated rules, conditions and considerations for the actual deployment of the EU FCT RDE. The ultimate ambition of LAGO is to go beyond the

¹ Paragraph 5 (b) of this Article provides for an exemption if such information proves impossible or would involve a disproportionate effort, for processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes. In this case, subject to the conditions and safeguards referred to in Article 89(1) GDPR, the controller shall take appropriate measures to protect the data subject's rights and freedoms and legitimate interests, including making the information publicly available.

creation of a common repository in order to innovate the FCT data-oriented research sphere by creation the crucial foundations for the sustainable, safe and trusted creation, co-creation, sharing and maintenance of training and testing datasets for the FCT research domain.

2. Data Controller

Data Controller: Centre for Research & Technology – Hellas (CERTH)/Information Technologies Institute (ITI), 6th km Harilaou - Thessaloniki, 57001, Thessaloniki - Thessaloniki, Greece.

Project Coordinator: ENGINEERING INGEGNERIA INFORMATICA SPA (ENG), established in PIAZZALE DELL'AGRICOLTURA 24, ROMA 00144, Italy

3. Data Processing

The purpose of data collection in this project is to extract useful information from public Web sources to discover and create datasets for the research community (i.e., Research Organizations, Academic Institutes). The collected datasets will be used in FCT research to train respective models.

With respect to the processing of personal data, the applicable legal ground for such processing activities is the legitimate interest of the data controller CERTH pursuant to Article 6(1)(f) GDPR. The processing is in fact necessary for scientific purposes and for the purposes of enhancing the understanding of the process of grooming, in accordance with recitals 49 and 50 GDPR.

What personal data is being processed?

The following categories of personal data publicly available will be processed with full respect of the terms and conditions of the relevant data provider (i.e., X):

- Social media posts (i.e., tweets), including the language, textual content and embedded multimedia content
- Social media account usernames

No special categories of personal data (art. 9(1) GDPR) are foreseen to be collected (at least not intentionally), nor data relating to criminal convictions (art. 10 GDPR)². In any case and in accordance with the data minimisation principle, only the parts of the social media posts that are deemed necessary for the project's objectives will be processed subject to a privacy-by-design technique, while the majority will be deleted immediately. All data will be collected in accordance with the licences and terms & conditions of the data providers. All data will be gathered only from public accounts, with the permission defined by X (former Twitter) and in compliance with the respective terms of use, including the ones referred explicitly to the terms of use on behalf of minors, thus in accordance with user expectation of privacy. All collected data will be

² In case any criminal activity is witnessed or uncovered during this research activity, the research team will be required to share this information and all necessary (pseudonymised) data with the appropriate and responsible authorities.

pseudonymised. The names of users that posted a tweet or are mentioned inside a tweet are all encrypted with a cryptographic cipher and replaced with alphanumeric characters. Data minimisation will also be applied, i.e., only data that are necessary for the purposes of the project will be processed. Further, details are provided in the “What is the purpose of the processing” section.

What is the purpose of the processing?

The above data will be required for the duration of the project: (i) for scientific research purposes, (ii) to facilitate the functionality of other modules of the project, and (iii) for demonstration purposes.

Data security

CERTH, as Consortium partner in the LAGO project, implements appropriate technical and organizational measures to ensure an appropriate level of protection against the risks arising from processing, such as accidental or unlawful destruction, loss, alteration, unauthorised disclosure, or access. The server hosting this database is accessible only by authorised users through authentication (using passwords of high complexity). A firewall will also be in place to allow only specific (whitelisted) Internet Protocol address (IPs) to access the server and to restrict the access of each whitelisted IP only to specific ports/services. Different access privileges to the data are available to ensure that the authorised users will only have access to the stored data on a need-to-know basis, i.e., to the stored pseudonymised data needed to fulfil their tasks. Devices that will store a backup of the data will follow the same security procedures as the main server. For any remote interactions with the server (e.g., remote control or data transfer), secure protocols such as ssh/sftp are used. Any processing of the data is performed on that server. In case processing will be needed on other machines, the same security measures of the server will be applied to the respective machine.

Will the collected data be shared?

The collected personal data may be disclosed: (1) to all partners of the Consortium, through a password protected system; and (2) if this is required to third parties (including data processors if exist) for the fulfilment of our legal obligations or is necessary for the fulfilment of the above data processing purposes and is in compliance with the applicable legal framework. It is also highlighted that no personal data will be transferred outside the European Union (EU) or the European Economic Area (EEA).

Who will be responsible for all of the data when this study is over?

When this study is over, CERTH will be the only one responsible for the information collected.

How long will data be stored?

The storage duration of the data in their anonymised or pseudonymised form will be the duration of the project plus five (5) years after the end of the project [i.e., October 2027], to be available for demonstration in case of an inspection or an audit, as long as required to achieve the above purposes of processing, unless a longer retention period is required by law or for the establishment, exercise or defense of legal claims

Will the collected personal data be used for other purposes?

All personal data will not be processed for any other purposes outside of those specified in this document.

Will the collected data be processed by automated tools supporting decision-making?

The collected data will not be processed by automated tools supporting decision-making. After hashing of any account information, the researchers will not be able to trace back the data back to the original owner.

What are your rights?

The data subject rights under GDPR are contained within articles 12-23 and 77. Some of your most important rights include:

- *Right to information:* you may request information about whether we hold personal information about you, and, if so, what that information is and why we are holding it. This information shall be provided within a reasonable period after obtaining the personal data, but at the latest within one month of receipt of the request.
- *Right to access:* you may request to receive a copy of the personal information we hold about you and to check that we are lawfully processing it.
- *Right to rectification:* you may ask us to rectify the information that we hold about you in case you consider that something is missing or is incorrect.
- *Right to erasure:* based on the grounds laid out in Art 17, you may ask us to erase your personal data at any given moment.
- *Right to object:* you may request us to stop processing their personal data based on the legal ground stated in Art 21 GDPR
- *Right to data portability:* you have the right to request the transfer of your personal data in an electronic and structured form to another party or directly to you. This enables you to take your data from us in an electronically usable format and to be able to transfer your data to another party in an electronically usable format.
- Lodge a complaint with the Hellenic Data Protection Authority (<https://www.dpa.gr>).

Please note that the aforementioned rights may be restricted in the light of the GDPR (e.g. art. 89 par. 2) and the applicable national data protection legislation.

For the exercise of the rights and for any other data-related information, data subjects may contact us at m4d_ethics@iti.gr