

# PathoCERT 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) GDPR<sup>1</sup>), 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 PathoCERT responsible for collection of data from online sources.

Data will be collected from:

- i. Public social media posts from Twitter, related to the quality of water areas (e.g., lakes, rivers) near urban regions and the clarity of drinking water.

### 1. The Project

Pathogen Contamination Emergency Response Technologies ([PathoCERT](#)) is a EU funded project aiming to strengthen the coordination capability of the first responders in handling waterborne pathogen contamination events, allowing the rapid and accurate detection of pathogens, improving their situational awareness and their ability to control and mitigate emergency situations involving waterborne pathogens. To achieve this objective, the project will research and demonstrate Pathogen Contamination Emergency Response Technologies (PathoCERT), a collection of novel, cost-effective and easy-to-use technologies, tools and guidelines, which will be field validated by the first responders.

### 2. Key Contacts

The Centre for Research & Technology – Hellas (6th km Harilaou - Themi, 57001, Themi-Thessaloniki, Greece) is the Data Controller, while the project coordinator is UCY.

### 3. Data Processing

The purpose of data collection in this project is to extract useful information by citizens observations through Twitter related to the quality of water areas (e.g., lakes, rivers) near urban regions and the clarity of drinking water. Further processing of the collected data will aim to

---

<sup>1</sup> 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.

estimate of their relevance to the examined use cases based on machine learning techniques to filter out irrelevant statuses, detect fake tweets, extract locations mentioned in the text in order to geo-localize them, detect any presence of nudity in the tweet's images and exclude tweets with inappropriate/pornographic text content. The legal basis under which the data shall be processed is article 6(1)(f) GDPR. The legitimate interest of the controllers lies with the pursuit of scientific research purposes confirmed by the European Commission.

#### **4. What personal data is being processed?**

The following categories of personal data publicly available in Twitter

- Language, tweet id, text, date, if the tweet is retweet or/and quote tweet, timestamp  
User name, user id, user's number of friends and followers, account age, tweet ratio (user's tweets per day), if the user is verified by Twitter

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). 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 the social media platforms (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 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.

#### **5. 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, especially in terms of their usefulness in the relevant use cases, and (iii) for demonstration purposes.

#### **6. Data security**

The PathoCERT 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. 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 the social media platforms and in compliance with the respective terms of use, thus in accordance with user expectation of privacy. 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. 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) 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., that the authorised users will have access only to the stored data needed to fulfil their tasks.

#### **7. 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. The information collected will be also used to contribute towards several journal and conference publications as well as scientific contests, in line with Twitter Policy. Non-EU countries (Republic of Korea) will be involved in PathoCERT research activities. In this context, personal data is foreseen to be exported for the needs of their research as described in the Proposal. Everything will be performed in compliance with Chapter V GDPR, as the EC has issued an adequacy decision for the Republic of Korea<sup>2</sup>.

#### **8. Who will be responsible for all of the data when this study is over?**

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

#### **How long will data be stored?**

The storage duration of the data will be the duration of the project plus five (5) years after the end of the project [i.e. 08/2023], to be available for demonstration in case of an inspection or an audit, or 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.

#### **9. Will the collected personal data be used for other purposes?**

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

#### **10. Will the collected data be processed by automated tools supporting decision-making?**

Your data will be used (i) for scientific research purposes, (ii) to facilitate the functionality of other modules of the project, and (iii) for demo purposes. Data collected from you will only be used to test the capabilities of the PathoCERT tools and you will not suffer any consequences of automated processing supporting decision-making.

#### **11. What are your rights?**

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

---

<sup>2</sup> [https://commission.europa.eu/system/files/2021-12/1\\_1\\_180366\\_dec\\_ade\\_kor\\_new\\_en.pdf](https://commission.europa.eu/system/files/2021-12/1_1_180366_dec_ade_kor_new_en.pdf)

- *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:* you may ask us to erase your personal data at any given moment without a specific reason.
- *Right to object:* you may request to stop processing, delete or remove your personal data at any desired moment where there is no good reason for us continuing to process it.
- *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 your rights and for any other data-related information you may contact us at [m4d\\_ethics@iti.gr](mailto:m4d_ethics@iti.gr)