Resilmesh

securing cyber infrastructures

Data Management Plan

Deliverable Number	**D1.6**
Deliverable Details: Data management plan describes how generated data will be managed.	
Deliverable Leading:	**JAMK**
Due Date:	**29/2/2024**
Submitted Date:	**29/2/2024**
Author(s)	**Vesa Vertainen, Tuomo Sipola**
Reviewer(s):	**Brian Lee (TUS), Jassim Happa (RHUL)**

Version History

Version	By	Date	Changes
0.1		22/11/2023	Draft
0.2		12/12/2023	Fixes
0.3		8/1/2023	Clarifications and elaborations, updated annex
0.9		12/1/2024	Moved to the Resilmesh template
A1		5/2/2024	Renamed according to project handbook practices
A2		26/2/2024	Corrections according to review
A3	VV	26/2/2024	New template
A4	TS	29/2/2024	Incorporate changes according to 2nd review
A	TS	29/2/2024	Release version

“Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or European Commission. Neither the European Union nor the granting authority can be held responsible for them.”

Table of Contents

Funded by the European Union

Introduction

This Data Management Plan (DMP) addresses all issues regarding the use of data within the consortium including the public release, posting, curation, and preservation of data during and after the project's lifetime. It follows the Findable, Accessible, Interoperable, Re-usable (FAIR) principles and is updated regularly to ensure appropriate data management and a high level of data quality and accessibility. This way we adhere to the principle “as open as possible, as closed as necessary”.

ResilMesh’s goal is to develop appropriate methods and tools to improve resilience capabilities and capacities in organisations by providing them with methods and tools to better:

- manage the complexity of their digital infrastructures and services,
- combat advanced persistent threats (APTs).

Project’s starting point is that “you can’t secure what you don’t understand”. Therefore, to address the above challenges ResilMesh will develop a cyber situational awareness (CSA) based security orchestration and analytics toolset to enable organisations achieve real time defence of essential business functions. Specifically, ResilMesh will help CyS organisations:

- Reduce CyS attack surface impact by developing tools to combat complexity (better visibility of assets and services and their dependencies), heterogeneity (interoperability and extensibility and dispersed infrastructure (flexible placement of security controls across the CyS infrastructure).
- Combat APT sophistication by developing advanced AI algorithms and tools for early and ongoing attack detection and prediction and improved situation and risk awareness.
- adapt to evolving security architectures and best practices by highlighting ResilMesh enabled security best practices to prepare for disruption by APTs (O3) as well as ResilMesh ‘zero trust ready’ approaches.

The purpose of this document is to provide an overview about ResilMesh’s Data Management Plan. The purpose of the plan is to ensure that there is a set of protocols and procedures supported by the project in place for handling of research, development, personal, and otherwise sensitive data.

The plan outlines how the consortium partners will treat data responsibly as we collect data to develop and improve ResilMesh, in addition to overviewing the data-handling philosophy for when ResilMesh is live and active after the project. Our data-handling and ethics are informed by activities by:

- **The European Union** – particularly in existing ethical and legal frameworks outlined in this document.
- **Academic ethics boards** – within the ResilMesh consortium, each academic institution have their own ethics committee. Each time an academic institution collects data that may have ethical concerns associated with it, the academics responsible for collecting the data put forward a case for their respective ethics

committee, outlining the data collection, processing, and storage issue to the board. The ethics committee then decides whether this is appropriate and whether any changes are necessary for the handling of data to be ethically sound.

- **The ResilMesh advisory board** – the project has an advisory board that includes ethics and data protection experts who continually contribute to critiquing the design and implementation of ResilMesh, in order to identify issues that emerge that the academic ethics board have been unable to identify.
- **Data Protection Authorities (DPAs)** – the consortium partners each reside in different nations. We also examine guidelines that they provide as part of our project.
- **An ethics research framework developed for this project**. This deliverable informs all other deliverables w.r.t. development and deployment of what to do (ethically and legally) through a peer-review approach to considering ethical, FAIR and legal challenges.

Overview of Ethical Frameworks in the EU

Research ethics and human rights are connected by overlapping and influencing each other. Several ethical frameworks exist in the EU. Ethics is an integral part of research, and therefore, ethical compliance is essential. The following discussion presents key ethical frameworks in the EU considered by the consortium.

- Charter of Fundamental Rights of the European Union[1]
- European Convention on Human Rights[2]

Within the EU Charter of Fundamental Rights, data protection is included under Article 8, stating that everyone has the right to the protection of personal data. Furthermore, the European Convention on Human Rights issued by the Council of Europe recognises the right to privacy under Article 8.

- The European Code of Conduct for Research Integrity[3]

The researchers and participating partners in the ResilMesh project adhere to the principles of the European Code of Conduct for Research Integrity of the European Science Foundation, including honesty, reliability, objectivity, independence and

[1] EU European Union, "Charter of Fundamental Rights of the European Union," Official Journal of the European Communities (2000/C 364/01), 2000. [Online]. Available: http://www.europarl.europa.eu/charter/pdf/text_en.pdf. [Accessed 2017]; European Parliament and Council, "Directive 95/46/EC of the European Parliament and of the Council, Official Journal L 281, 23.11.1995, pp. 31-50, Luxembourg," 1995. [Online]. Available: http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:31995L0046&from=EN. [Accessed 2024].

[2] Council of Europe, "Convention for the Protection of Human Rights and Fundamental Freedoms," 2010. [Online]. Available: http://www.echr.coe.int/Documents/Convention_ENG.pdf. [Accessed 2024].

[3] ALLEA, "The European Code of Conduct for Research Integrity – Revised Edition 2023," 2023. [Online]. Available: http://www.doi.org/10.26356/ECOC. [Accessed 2024].

impartiality. The principles also cover duty of care, fairness and responsibility for future generations.

- Horizon Europe Rules for Participation (Regulation No 2021/69)[4]

Regarding personal right (paragraph 71), classified information (Article 20), ethics (Article 19) and exploitation and dissemination (Article 39).

- Association of Internet Researchers guidelines[5]

These ethical guidelines present another point of view to ethical research.

Data Protection and Privacy in the EU

Data Protection Directive (95/46/EC)

The EU's Data Protection Directive[6] aims to harmonise national laws on privacy and data protection. It defines which data processing is legitimate,[7] focusing on operations performed upon personal data. The directive defines, under Article 2, personal data, processing of personal data, personal data filing systems, controllers, processors, third parties, recipients and data subject's consent.

Directive on Privacy and Electronic Communications (ePrivacy Directive) (2002/58/EC)

The 2002 ePrivacy Directive[8] extends the Data Protection Directive. It considers furhter cookies, spam and confidentiality of communications.

[4] European Parliament and Council, "Regulation (EU) 2021/69 of the European Parliament and of the Council," Official Journal L 170, 15/5/2021, pp. 1-68, Strasbourg: European Commission, 2021. [Online]. Available: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32021R0695&from=EN. [Accessed 2024].

[5] A. Markham, E. Buchanan and A. E. W. Committee, "Ethical Decision-Making and Internet Research: Recommendations from the AoIR Ethics Working Committee (Version 2.0)," Association of Internet Researchers, 2012. [Online]. Available: https://aoir.org/reports/ethics2.pdf. [Accessed 2024].

[6] European Parliament and Council, "Directive 95/46/EC of the European Parliament and of the Council," Official Journal L 281, 23.11.1995, pp. 31-50, Luxembourg: European Commission, 1995. [Online]. Available: http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:31995L0046&from=EN.

[7] A. Richter, "The Protection on Privacy and Personal Data on the Internet and Online Media," Committee on Culture, Science and Education, Parliamentary Assembly of the Council of Europe, 2011. [Online]. Available: http://www.assembly.coe.int/CommitteeDocs/2011/RihterviepriveeE.pdf. [Accessed 2024].

[8] European Parliament and Council, "Directive 2002/58/EC of the European Parliament and of the Council," Official Journal L 201 , 31/07/2002, pp. 37-47, Brussels: European Commission, 2002. [Online]. Available: http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32002L0058&from=en. [Accessed 2024].

Data Retention Directive (2006/24/EC)

The Data Retention Directive[9] demands telecommunications providers within the EU to retain customers' traffic from six months to two years from the starting date of the communication. However, later the European Court of Justice of the European Union declared the directive invalid, as the rights to privacy and personal data protection were endangered.[10]

General Data Protection Directive (2016/679)

The General Data Protection Directive[11] (GDPR) has been in effect since 2018. It includes main definitions related to processing of personal data. **Personal data** is any information that relates to an identified or identifiable natural person. Such data may operated upon bu **processing**. Other essential definitions include restriction of processing, profiling, pseudonymisation, filing system, controller, processor, recipient, third party, consent, personal data breach, genetic data, biometric data, data concerning health, main establishment, representative, enterprise, group of undetakings, binding corporate rules, supervisory duty, supervision authority concerned, cross-border processing, relevant and reasoned objection, information society service and international organisation.

GDPR Roadmap for ResilMesh

This section describes a roadmap for ResilMesh towards GDPR compliance:

A. Scope: In order to fall under the scope of the GDPR, the case under investigation must fulfil all the following conditions:

1. the data under scrutiny fall within the ambit of the definition personal data (art. 4 § 1), and
2. processing is conducted by automated means or by means other than automated as part of a filing system (art. 2 § 1), and
3. the case falls within the material scope of the GDPR (art. 2 § 2-3), and
4. the case falls within the territorial scope of the GDPR (art. 3).

B. Lawfulness of Processing: If the case under investigation falls under the scope of the GDPR, the next step of legal evaluation is the lawfulness of processing. Data

[9] European Parliament and Council, "Directive 2006/24/EC of the European Parliament and of the Council," Official Journal 105, 15/03/2006, pp. 54-63, Luxembourg: European Commission, 2006. [Online]. Available: http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2006:105:0054:0063:EN:PDF. [Accessed 2024].

[10] Court of Justice of the European Union, "The Court of Justice Declares the Data Retention Directive to Be Invalid," Press release No 54/14 Judgement in Joint Cases C-293/12 and C-594/12 Digital Rights Ireland and Seitlinger and Others, Luxembourg, 8 April 2014, 2014. [Online]. Available: http://curia.europa.eu/jcms/upload/docs/application/pdf/2014-04/cp140054en.pdf. [Accessed 2024].

[11] European Parliament and Council, "Regulation (EU) No 2016/679 of the European Parliament and of the Council," Official Journal L 119, 27 April 2016, pp. 1-88, Brussels: European Commission, 2016. [Online]. Available: http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R0679&from=en. [Accessed 2024].

Funded by the European Union

processing is considered lawful, if one of the following legal bases is fulfilled (arts. 6 § 1 and 9 § 1):

1. consent, or
2. contractual necessity, or
3. legal obligation, or
4. vital interests, or
5. public interest, or
6. legitimate interests, or
7. special categories of processing.

C. Data Protection Principles: If data processing in the case under investigation is lawful, the next step of legal evaluation is the compliance of data processing with the general principles of the GDPR (art. 5 § 1):

1. fairness, lawfulness and transparency, and
2. purpose limitation, and
3. minimisation [Proportionality], and
4. accuracy, and
5. storage limitation, and
6. integrity and confidentiality.

D. Data Controller / Processor Obligations: If data processing in the case under investigation complies with the general principles of the GDPR, the next step of legal evaluation is the compliance of data controllers / processors with certain obligations under the GDPR, as follows:

1. respect of data subjects' rights (arts. 12-20), and
2. implementation of technical and organisational measures (art. 24), and
3. implementation of confidentiality and security measures (art. 32).

Accountability Requirements. If data controllers / processors comply with the foregoing obligations under the GDPR, the next step of legal evaluation is the fulfilment by data controllers / processors of the following accountability criteria before the Data Protection Authorities (DPAs):

1. data protection by design and by default (art. 25), and
2. appointment of representative[s] before regulators for controllers / processors not established in the EU (art. 25), and
3. record keeping (art. 30), and
4. cooperation with DPAs (art. 31), and
5. data breach notification to DPAs (art. 33), and
6. data breach notification to data subject[s] (art. 34), and
7. data protection impact assessment (art. 35), and
8. prior consultation with DPAs (art. 36), and
9. mandatory data protection officers (art. 37).

It will be necessary specify high-level (human-level) protocols/procedures to address automation or human errors in data sharing or a guidelines for what organisations should do in the case of sensitive data leakage (intentional or not), the ResilMesh tool itself attacked etc.

Funded by
the European Union

We will describe a number of scenarios and planned responses closer to the pilot studies (to be included in the pilot design documents). Our starting point will be to use this deliverable and the GDPR as our foundation for actions to take – e.g. reporting any data leaks or attacks to the data protection authority (DPA) in a timely manner and responsibly disclose any attacks to affected parties.

Data Summary

All project partners provide input on the data types they collect and store, the protective measures taken and follow the stipulated principles and guidelines of the DMP. Specific attention is given to requirements gathering and analysis activities, which involve acquiring **questionnaire**, **interview**, **observational study**, and **user study data**, to ensure that these adhere to legal and ethical guidelines outlined in the DMP and that users are transparently informed and fully consent to these methods. The ResilMesh solution itself will involve **cyber threat intelligence exchange** between partners. Such data considered can constitute a sensitive/personal information and must be treated appropriately to address the data privacy issues.

Other types of data to be collected in the project are, for example, the project management data, source code and datasets for machine learning models. Project management data includes progress reports, financial reports, communication, and dissemination data etc.

A data collection document will be created as a living document, which will be based, e.g., on what is presented in Annex 1.

The machine learning models developed in the project, will re-use publicly available datasets and are also trained with datasets generated within the project. Datasets may be generated (i) as part of WP4 in the form of threat awareness data, (ii) as part of WP5 in the form of situational data and (iii) as part of the cyber range activities in WP6 also using data generation tools from use-case partners ALWA and ALIAS. The datasets are used in developing AI anomaly detection algorithms to detect suspicious events and attacks at both edge and cloud for host and network data including logs and emitted events. Datasets are also needed to train models to correlate security events or raw data with the goal to predict attack evolution, determine the root causes of attacks or anomalies, and reduce the number of events, e.g., false positives, submitted to a SOC operator. In addition to datasets, there are other research outputs that might be useful to other parties outside the project as well (SOC operators etc.). These are MISP artifacts, playbooks, machine/deep learning models and so on. An overview of possible data types, and their expected sizes, is shown in Table 1.

Table 1: Overview of data collected and generated.

Content	Data type	Storage/system	Expected size

Funded by the European Union

Datasets for deep learning models	Datasets (CSV and other formats)	Github	100-1000 Gigabytes
Project administrative material, Course of Action (CoA) playbooks, Dissemination material…	Text documents	Google Docs	1-10 Gigabytes
Machine learning models	Source code (Python)	Github	100-500 Megabytes
Documents, images, spreadsheet, deliverables…	Administrative documents	Google Drive	1-5 Gigabytes
Project resource planning and financial monitoring data	Administrative database	Emdesk	100-500 Megabytes
Dissemination material	Presentations	Google Drive, Local PCs	1-10 Gigabytes
MISP artifacts	JSON	Gitlab / MISP threat sharing communities across EU	1-100 Gigabytes
Project communication	Online chat	Slack	-

Template for Research Data

A provisional template to identify and describe the research data collected and generated in the project, is included at the end of this document's **Annex 1**. The table currently has sample answers that ought to be replaced with correct ones. The table has the following fields:

- **Related WP**: The work package where the data is collected or generated in.
- **Unique ID:** A unique identifier for the dataset (using a common prefix "RESILMESH-").
- **Description:** A description of the type of data involved.
- **How is the data gathered:** A description of how the data was collected or produced, i.e., specialist instruments or tools.
- **Origin/owners/users:** The origin, owners, or users of the data
- **File format:** Data format, for example PY, JSON, CSV
- **Data storage:** Where and how is the data stored.
- **Required disk space:** Expected size of data in question.
- **Is the data confidential:** Is the data confidential, i.e., business secrets.
- **Does it contain personal data:** Does the data contain personal information.
- **Does it contain sensitive data:** Does the data contain sensitive information.
- **License**: If reused data, license that allows you to use it (and to share?).
- **Responsibility for data management:** Which participant is responsible for it.

- **Applicable legislation:** Applicable legislation on data protection, which legislation?
- **Sharing/access type for reuse:** How can the data be accessed.
- **Collected/produced/reused:** Collected for this project OR produced as an outcome OR previously collected data reused?
- **Pilot/country:** Pilot participant involved.

FAIR DATA

ResilMesh complies with the FAIR principles (findable, accessible, interoperable, and re-usable). Research outputs are scientific publications, data, software, algorithms, protocols, models, workflows and other engineered results and processes that resulted from the project. The research outputs are organized and structured to be easy to access, understand and reuse. As far as possible, the data is optimized to be machine-accessible without human intervention or with minimal human intervention.

Making data findable, including provisions for metadata

The project results will be made available via the project website, open-access platforms, and channels. Contributions to the data repositories will include persistent and unique identifiers (PID) and metadata. Search keywords are provided in the metadata for optimal discovery and potential re-use.

Common file formats, standards, and well-defined practices will be used whenever possible to ensure machine-readability, interoperability, and reusability.

As the data gathered during the pilots may contain information specific to a given organization, a data sharing agreement will be made with the pilot participants. All partners will ensure that their own national data protection standards are applied to all data they gather or process during the action, as explicitly mentioned in the Consortium Agreement.

The data is registered on a repository as soon as it is created, and metadata is added to make it “findable online by giving the name of the dataset, authors, date of creation, DOI, etc”. At this stage at least a README file is deposited publicly available.

Making data accessible

Repository: The data is deposited in trusted repositories and made open as soon as possible. An open licence, CC-BY (or CC0) licence, is used. All public deliverables will be available on the project website and will be given Persistent Identifiers (PIDs).

Data: This project fully subscribes to the principles of Open Science as defined by the European Commission and already actively rolled out by the research partners of the Consortium. Scientific publications will be offered under an OA license, the copyright retainment clause specified by the Grant Agreement, will be respected, so that Creative Commons Attribution International Public License (CC-BY) or equivalent will be sought to the maximum extent possible.

Metadata: Following the "as open as possible, as closed as necessary" principle, the data could be closed, if necessary, but the metadata is FAIR and has a CC0 licence.

Making data interoperableCommonly used open-source software is used within the project. Documents, source code, MISP artifacts etc. are written in formats that are readable with open-source software. The project also follows open standards for security interoperability such as those from the Open Cybersecurity Alliance[12]

Increase data re-use

The results of the project are shared in commonly used open-source formats, making it easy to re-use. Information about the tools needed to re-use or examine the data, as well as other information about research outputs, is provided. These outputs can be, for example, datasets, source code, algorithms, models, workflows, playbooks and so on. ResilMesh follows an open-source strategy for implementation tasks, reusing and extending open-source components and technologies, such as Kubernetes, ISTIO, and Apache Airflow.

Other research outputs

Software, software documentation and AI models will use open licensing where possible. Restrictions to licensing and availability may come from partner background, base software licensing, grant agreement and confidentiality levels of data, models, and research outputs and results.

Allocation of resources

Costs for making data or other research outputs FAIR are considered as follows:

Publications

Scientific publications will be offered under an open access (OA) license. The choice between *Open Research Europe*, an OA Journal, or a *Subscription Journal* with retention of rights will be made on an ad-hoc basis, depending on whether suitable, reputable OA journals exist for a given topic, and how these compare to open access options offered by reputed OA Journals in the field. In any case, the copyright retainment

[12] https://opencybersecurityalliance.org/

clause specified by the Grant Agreement, will be respected, so that Creative Commons Attribution International Public License (CC-BY) or equivalent will be sought to the maximum extent possible.

Data, software & notebooks

The public data, software and notebooks that are generated in the project, can be saved to free public platforms, such as Github.com. Project management data, and other non-public data is saved on free platforms, such as Google Drive, or other (internal) platforms in use by each consortium party.

The possible costs related to research data/output management are eligible as part of the Horizon Europe grant. Table 2 provides the data processes and their responsible roles.

Table 2: Data processes and responsible entities.

Description	**Responsible entity**	**Responsible person**
DMP creation and updating	Project data manager	Tuomo Sipola (Jamk), Vesa Vertainen (Jamk)
Collection and curation of data in pilots	WP8 leader	Jassim Happa
Processing and preservation of data	WP leaders	WP Leaders
Publishing and sharing data	Data owners	All

Data security

In the processing of all personal data the consortium will comply with the Data Protection principles which are set out in General Data Protection Regulation (GDPR). Data security relies on the security of the various service providers. Potential service providers and their GDPR compliancy is listed in Table 3 below.

Table 3: Service providers and GDPR compliancy.

Service	**GDPR Compliant**	**Statements about data location, and links to policy**
Emdesk	Yes	“Data primarily within the European Economic Area. However, we have service providers and operations in several geographical locations. As such, we and our service providers may transfer your personal data to, or access it in, jurisdictions outside the European Economic Area." https://www.emdesk.com/privacy-policy
Google Drive	Yes	“As an administrator, you can store your covered data in a specific geographic location by using a data region policy.” https://www.google.com/intl/en-GB/about/company/user-consent-policy-help

		https://cloud.google.com/privacy/gdpr
Slack	Yes	"Data residency for Slack allows global teams to choose the region where certain types of data at rest are stored. If you're not using data residency, your data will be stored in the US (AWS)." https://slack.com/intl/en-gb/trust/privacy/privacy-faq https://slack.com/trust/compliance/gdpr https://slack.com/intl/en-ie/help/articles/360035633934-Data-residency-for-Slack
OneDrive	Yes	"You can control where data resides on a granular level, specifically, on a per-user basis." https://www.microsoft.com/en-us/trust-center/privacy/gdpr-overview https://techcommunity.microsoft.com/t5/microsoft-onedrive-blog/gdpr-compliancy-with-onedrive-and-sharepoint/ba-p/191126
Github.com	Yes	"We transfer personal data from the European Union …to other countries, some of which have not yet been determined by the European Commission to have an adequate level of data protection …we use a variety of legal mechanisms, including contracts, such as the standard contractual clauses published by the European Commission under Commission Implementing Decision 2021/914, to help protect your rights and enable these protections to travel with your data." https://docs.github.com/en/site-policy/privacy-policies/github-privacy-statement#european-data-protection-rights-notice https://github.blog/2018-04-19-updates-to-our-privacy-statement-and-terms-of-service
Gitlab.com	Yes	Hosted on Google Cloud. https://about.gitlab.com/privacy/privacy-compliance/

Ethics

All project partners will follow the ethical guidelines as defined in this data management plan. These guidelines will ensure privacy and security in all activities that involve participation of end users. We will develop **procedures and templates** (e.g., informed consents, information sheets) for the collection, use and storage of personal data that will be acquired through focus groups, surveys, user evaluations and interviews. For activities that require specific ethical approval, this ethical approval will be sought prior to each activity involving users or their data according to the procedures of the party performing the activity.

The use of Artificial Intelligence

ResilMesh will use AI to develop several algorithms and asserts the following measures will be taken to ensure the algorithms are developed and used in a safe, secure, and ethical manner. Following the ALTAI guidelines, we identify three areas to address:

Human Autonomy and Oversight: In certain cases, AI-based correlation may optionally be used to reduce SOC analyst workload by filtering alarms with input from the analyst. This operational mode is Human-in-Command. The analyst is fully aware of interaction with AI and may switch off the tool at any time. Training will be provided for use-cases where the tool is used.

Technical robustness: Deep learning models and datasets are susceptible to a variety of adversarial attacks and need to be carefully trained to be resilient to attack. ResilMesh will conduct an analysis at the beginning of the project to better understand the types of adversarial attacks that might be germane to the use case environments and will consequently determine and apply the appropriate techniques to counter these attacks. ResilMesh also provides a fallback plan in all cases where AI is used. AI based anomaly detection may fall back to rule-based or statistical based techniques if required though the use of IDS and SIEM systems while statistical techniques such as time series or data mining may be used for attack prediction instead of deep learning approaches. In the case of cyber forensics AI techniques may be set aside and manual approaches used instead. Accuracy is achieved using well-formed development and evaluation processes. Moreover, in several cases we propose to use a number of algorithms in parallel to achieve better robustness in the accuracy e.g., through the use of ensemble models for anomaly detection decision fusion or the dual use of time-series and deep learning for prediction. Reliability will be achieved using well-formed development and evaluation processes and carefully selected data sets to ensure general and reproducible results.

Privacy and Data Governance: No personal data is used for training AI systems in ResilMesh. Datasets are either well-known anomaly datasets or will be generated in the project. Where federated learning may be used as in the civic infrastructure use case, strong cryptographic techniques will be used to achieve data privacy when sharing model parameters. Moreover, we will leverage these techniques to address the call issue of mass surveillance and privacy of personal spaces using federated learning in anomaly detection personal space, such as mobile phones or home networking, in the open call use cases.

The full Assessment List for Trustworthy Artificial Intelligence (ALTAI) for self-assessment can be downloaded from the European Commission page: https://digital-strategy.ec.europa.eu/en/library/assessment-list-trustworthy-artificial-intelligence-altai-self-assessment

Ethics issues checklist for the use of Artificial Intelligence

Parties that use AI may use the following checklist (Table 4) to evaluate their activities.

Table 4: Checklist for ethics issues.

	Yes/ No	**Information to be provided**	**Documents to be provided / kept on file**
Does this activity involve the development, deployment and/or use of Artificial Intelligence-based systems?		Explanation as to how the participants and/or end-users will be informed about: - their interaction with an AI system/technology - the abilities, limitations, risks, and benefits of the proposed AI system/technique - the way decisions are taken and the logic behind them. Details on the measures taken to avoid bias in input data and algorithm design. Explanation as to how the respect to fundamental human rights and freedoms (e.g., human autonomy, privacy, and data protection) will be ensured. Detailed explanation on the potential ethics risks and the risk mitigation measures.	Detailed risk assessment accompanied by a risk mitigation plan (if relevant). These must cover the development, deployment, and post-deployment phases. Copies of ethics approvals (if relevant)
Could the AI based system/technique potentially stigmatise or discriminate against people?		Detailed explanation of the measures set in place to avoid potential bias, discrimination, and stigmatisation.	
Does the AI system/technique interact, replace, or influence human decision-making processes?		Detailed explanation on how humans will maintain meaningful control over the most important aspects of the decision-making process. Explanation on how the presence/role of the AI will be made clear and explicit to the affected individuals.	Information sheets/Template Informed consent forms (if relevant).
Does the AI system/technique have the potential to lead to negative social impact?		Justification of the need for developing/using this technology. Assessment of the ethics risks and detailed description of the measures set in place to mitigate the potential negative impacts during the research, development, deployment, and post-deployment phase.	For serious and/or complex cases: Algorithmic impact assessment/human right assessment. These must cover the development, deployment, and post-deployment phases.
Does the AI to be developed/used in the project raise any other ethical issues not covered by the questions above?		Detailed explanation on how the potential ethics issues will be addressed, and the measures set in place to mitigate ethics risks.	Detailed risk assessment accompanied by a risk mitigation plan. These must cover the development, deployment, and post-deployment phases.

Checklist adapted from: https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/common/guidance/how-to-complete-your-ethics-self-assessment_en.pdf

Data and Ethics Management Plan

The ResilMesh consortium aims to ensure the compliance of the performed activities with national and EU legislation, and with the basic ethical principles that represent the shared values upon which the EU is founded and that are laid down in the European Charter of Fundamental Human Rights. In this Section, we outline in-depth our Data and Ethics Management Plan.

Ethics Considerations for the ResilMesh Project

An important aspect of carrying out research in ResilMesh is to do so in a way that is ethical and respects privacy, dignity, and welfare of research participants. The project ensures that all our institutions have comparable ethical standards that researchers are required to meet. In addition, we consider the ethical requirements set out by the relevant competent authorities, including DPAs and ethics committees, in the countries where research is undertaken.

Ethical Code of Conduct in the ResilMesh Project

ResilMesh has set out a system of privacy and ethics governance that involves:

- An external advisory board that includes ethics and data protection experts.
- A number of project partners have their own ethics committees that they are responsible to.
- The ResilMesh project is developing a set of tool protocols: a human-level set of common practice guidelines for use of ResilMesh and rules that organisations connected to ResilMesh should abide by. This will be created for the Pilot Design.

Ethics review and legal compliance for the ResilMesh Project

The empirical studies conducted for the ResilMesh project have been reviewed by, and received ethics clearance through, Royal Holloway and Bedford New College's (RHUL) Research Ethics Committee. To protect the privacy of individuals, and to prevent unauthorised access to data, the project consortium is implementing strict procedures to safeguard the privacy of the participating individuals and is implementing information sharing policies that prevent unauthorised access to data that is available within the consortium. All research activities of the ResilMesh project that are carried under the EU research funding programme H2020 comply with the national, EU and international ethical and legal framework introduced above, such as the Charter of Fundamental Rights of the European Union and the European Convention on Human Rights. The project is working in compliance to the EU Directive 95/46/EC, the EU Directive 2002/58/EC, and the EU Regulation 2016/679.

Data-handling in the ResilMesh project

The ResilMesh project collects two types of data sources:

a. Empirical research data for the gathering of the ResilMesh requirements
b. Data collected via or created from the Pilots.

Empirical Research Data for the Gathering of the Project Requirements

For the purpose of gathering the ResilMesh design requirements, as well as testing them, we are conducting qualitative research that is collecting data through survey questionnaires, interviews, and observations of the participating organisations.

Anonymity and confidentiality: Data are anonymised to protect the confidentiality of the research participants. In particular, direct identifiers (e.g. names, addresses, etc.) and indirect identifiers (e.g. role within organisation) are removed from the data.

Data-handling and transfer: We store all collected research data (e.g. interview recordings and transcriptions, survey responses) securely by using hard-drive encryption. We transfer collected data only via using secure channels of communication (e.g. encrypted network connections, formatted and encrypted USB flash drives).

Use of data: The research subjects have the option to ask us to access the collected data. Furthermore, they can at any given point ask us to withdraw any of the collected data or information within the data that relate to them. The information collected under this project will be controlled via a protocol hierarchy for access: the partners of the project will have access only to processed data that have been fully anonymised.

Informed consent: We send a copy of the Participant Information Sheet to the research participants in advance. During the research itself, we provide another copy of the Participant Information Sheet and explain before why we are conducting this research, how it is being used for and discuss ethical issues. After conducting this research, we ask them to read and sign the Participant Consent Form.

Legal requirements: The project works in accordance with the ethical and legal requirements established by the European Commission and national authorities in the related project areas, in particular concerning data protection and privacy issue.

Sensitive information from the Use Case Partners: Questions (questionnaire and interviews) are reviewed and approved by ResilMesh members and their constituencies. If any issues arise with the questions we ask them to make suggestions for changes.

Pilots Data Management

A separate section on data management will be provided for the pilot design. This is because the details of the management will, by need, come from the completed design of the detailed piloting activities.

Annex 1 - Research Data Table

Use the table to identify and describe research data. Add/delete rows as needed. The table will be attached to the ResilMesh data management plan. This table is a modified version of the” Research Data Info Table” by the Aalto University and is licensed under a Creative Commons Attribution 4.0 International License.

WP	ID	Description	How is the data gathered?	Origin/ owners /users	File format	Data storage	Required disk space	Is the data confidential?	Contains personal data?	Contains sensitive data?	License	Responsible for data management	Applicable legislation	Sharing/ access type for reuse	Collected/ produced/ reused	Pilot/ Country
1	1234	Interview data	Audio recording		MP3	OneDrive	2GB raw audio, 150kb text	No	Yes. Direct identifiers like voice	Yes (health information)	N/A		GDPR		Collected	
		Machine learning models	Python code		.py	Local gitlab server	50MB	No	No	No	CC BY 4.0				Produced	
		Dataset “X”	Downloaded from https://www.unb.ca/cic/datasets/		CSV	Local gitlab server	7.5GB	No	No	No	CC BY 4.0				Reused	