Finansuoja Europos Sąjunga
NextGenerationEU

Mykolo Romerio universitetas

NAUJOS KARTOS LIETUVA

Project „Misijomis grįstų mokslo ir inovacijų programų įgyvendinimas“ (Project Nr. 02-002-P-0001) report

Project name: ***EPMwDC***

Responsible/Implementation partner (-s): ***MRU***

Action result: ***Modeling According to the Principles of Privacy by Design and Privacy by Default: System Analysis Requirements***

Finansuoja Europos Sąjunga NextGenerationEU

Mykolo Romerio universitetas

NAUJOS KARTOS LIETUVA

Modeling According to the Principles of Privacy by Design and Privacy by Default: System Analysis Requirements

Table of Contents

Introduction

The project “Research and development of an Espionage Prevention Monitor with Detection Capabilities (EPMwDC)” seeks to develop a unique product consisting of hardware and software modules, which would enable government and private organisations to prevent sensitive data leaks when a picture is taken of the screen, displaying confidential data, detecting and reporting to authorities if such incident is realised.

To comprehensively fulfill the requirements of Privacy by Design, Security by Design, and the broader regulatory landscape (including the GDPR, EU AI Act, NIS2, and NATO standards), system analysis and modeling must embed legal and technical safeguards into the system's core architecture before deployment.

Below is an in-depth elaboration of the key stages and specific technical requirements that must be accounted for during system analysis.

1. Detailed Stages of System Analysis and Implementation

1. Stage 1: Strategic Segmentation (Design Phase)

During the initial design phase, the system architecture must be strategically segmented to navigate the "Dual-Use Regulatory Trap". Developers must clearly define the intended purpose and separate the product lines (e.g., a Civilian version and a Military version) because they are subject to vastly different compliance tracks.

- **Civilian Systems:** Must integrate stringent "Privacy by Design" mechanisms, undergo conformity assessments, and comply with the EU AI Act, GDPR, and Cyber Resilience Act.
- **Military Systems:** While exempt from the EU AI Act, these systems must integrate "Security by Design" and align with NATO’s 6 Principles of Responsible Use (PRUs), focusing on traceability and interoperability within secure defense networks.

2. Stage 2: Data Governance and Training

System analysis must establish strict data governance protocols for the datasets used to train the AI model.

- **Data Quality and Provenance:** Training must utilize high-quality, unbiased datasets. For defense applications, this requires creating a "DARB-Ready File" (Data and AI Review Board) to document data provenance and prove the system was not trained on compromised data.
- **Bias Mitigation:** The system's computer vision must be modeled to ensure it does not misidentify specific skin tones or benign behaviors as threats, satisfying both GDPR fairness principles and NATO’s Bias Mitigation requirements.

3. Stage 3: Testing and Validation (Security & Robustness)

Before deployment, system analysis must mandate rigorous testing to ensure resilience against adversarial manipulation and environmental challenges.

- **Adversarial Testing (Red Teaming):** The AI models must be tested against "data poisoning" and model evasion (e.g., tricking the AI using anti-IR filters or specific camera angles) to meet the robustness requirements of the AI Act (Art. 15) and NATO reliability standards.
- **Sandboxing:** System validation should utilize regulatory sandboxes (such as the AI Act Regulatory Sandbox or NATO DIANA) to test compliance and performance in realistic or denied environments without facing immediate regulatory penalties.

4. Stage 4: Impact Assessments (DPIA & FRIA)

System analysis cannot proceed without legally mandated impact assessments conducted prior to deployment.

- **Data Protection Impact Assessment (DPIA):** Mandatory under GDPR Article 35 for systems that systematically monitor individuals (e.g., employees). The DPIA must evaluate risks, proportionality, and necessary privacy safeguards, and must be updated if the deployment context changes.
- **Fundamental Rights Impact Assessment (FRIA):** Required for high-risk deployers under the EU AI Act to verify that the system does not unduly infringe on civil liberties or employee privacy rights before the system goes live.

2. Specific Technical Requirements and Key Aspects

1. Data Minimization and Edge Computing

System modeling must guarantee that the technology operates without unnecessarily collecting personal data.

- **Real-time Overwriting:** The system should be engineered so that continuous monitoring or recording does not take place; video buffers must be instantly overwritten in real-time.
- **Trigger-based "Freezing":** Data must only be permanently saved ("frozen") when a positive threat detection (e.g., a camera lens) is confirmed. The system must effectively distinguish between a "threat" and a "non-threat" (like a compliant human operator) and avoid retaining data on the latter.
- **Edge AI Processing:** System analysis should mandate local processing directly on the hardware module. Raw images must be discarded immediately at the edge, and only necessary metadata (such as the timestamp, location, and the AI's confidence level) or a single alert-related frame should be transmitted to central servers.

2. State-of-the-Art Encryption and Pseudonymization

The design must embed a comprehensive encryption architecture from inception to protect data at rest and in transit.

- **Transmission Standards:** All communications to remote monitoring servers must utilize end-to-end encryption protocols, such as AES-256. Defense deployments may further require quantum-resistant algorithms or state-approved cryptographic modules (e.g., compatibility with Lithuania's *Saugusis tinklas*).
- **Pseudonymization:** To minimize privacy impact, the system should transmit pseudonymized alerts (e.g., flagging "User ID 12345" instead of an employee's name or sending a blurred image), with the re-identification keys held securely and separately.

3. Strict Access Controls, Audit Logging, and Cyber Resilience

Remote server architectures must be fortified against unauthorized access and maintain full traceability.

- **Role-Based Access Control (RBAC):** Access must be restricted based on the principle of least privilege, requiring multi-factor authentication (MFA).
- **Immutable Audit Logs:** The system must generate immutable logs tracking all system access, data processing activities, and user actions (e.g., logging an operator's review of an alert). Depending on national transposition (like Lithuania's cybersecurity laws), logs may need to be retained for at least 90 days.

- **Cyber Resilience by Design:** The hardware and software modules must feature secure boot mechanisms, hardened firmware, and automated over-the-air (OTA) update capabilities. Furthermore, developers must maintain a Software Bill of Materials (SBOM) tracking all third-party components to comply with the Cyber Resilience Act.

4. Human Oversight (Human-in-the-Loop)

System workflows must be legally and technically modeled to prevent automated decision-making from producing binding legal or disciplinary effects without human intervention.

- **Review Workflows:** If the AI detects a potential threat, it must flag the event (providing the annotated image and metadata) for an authorized security officer or analyst to review.
- **Reversible Outcomes:** The human operator must have the authority to classify the alert as a "false positive" or a "false negative," providing feedback that can reverse the automated outcome and help retrain the model.

5. Algorithmic Transparency and Explainability

System architecture cannot rely on a "black box" approach; it must provide meaningful information regarding how decisions are made.

- **Right to Explanation:** Following CJEU rulings (e.g., *Dun & Bradstreet*), companies cannot use trade secrets as a blanket refusal to explain the AI's logic to an affected data subject. The system must record explainability logs detailing the parameters that led to a "threat" classification.
- **Internal Transparency:** System documentation must be robust, and organizations deploying the system must provide transparent internal privacy notices to employees, clearly outlining the conditions under which monitoring and image capture might occur.

6. Automated Deletion and Pre-defined Retention Policies

Finally, retention periods must be hard-coded into the design phase, not left as a post-deployment afterthought.

- **Storage Limitation:** Incident images and metadata must only be stored for the strict duration required to investigate the event.
- **Automated Purging:** While the exact retention schedule is defined by the end-user entity's internal policies, the system must technically support and enforce automated deletion mechanisms once the defined retention period expires.

Annexes: System Analysis Requirements (According to Privacy and Security by Design)

Annex 1. Key Stages in System Analysis

Key Stages in System Analysis	
Strategic Segmentation (Design Phase)	• Defining the intended purpose of the system and separating the product lines (Civilian vs. Defense) to navigate different regulatory tracks.
Data Governance & Training	• Establishing strict protocols for the selection, annotation, and preparation of training data to ensure it is unbiased, relevant, and secure.
Impact Assessments (DPIA & FRIA)	• Conducting mandatory Data Protection Impact Assessments (DPIA) and Fundamental Rights Impact Assessments (FRIA) prior to deployment to evaluate proportionality and risks.
Testing & Validation	• Utilizing regulatory sandboxes and adversarial testing (Red Teaming) to prove the AI model's robustness and compliance before finalizing the analysis.

Annex 2. Compliance Checklist

Compliance Checklist	**[ ]**	• Determine whether the system is classified as "High-Risk" under the EU AI Act or falls under military exemptions.
	[ ]	• Ensure "Data Minimization by Default" is integrated into the analysis, planning for local processing (Edge AI) where images are immediately discarded.
	[ ]	• Plan for Algorithmic Transparency to satisfy the "Right to Explanation", ensuring the system can output logic regarding why a threat was flagged.
	[ ]	• For defense applications, create a "DARB-Ready File" (Data and AI Review Board) documenting data provenance and alignment with NATO's 6 Principles of Responsible Use.
	[ ]	• Ensure automated decision-making procedures do not bypass human review, avoiding prohibited AI practices.

Annex 3. Challenge Overcoming

Challenges	Suggestions to Overcome Challenges

• **The "Dual-Use" Regulatory Trap:** Attempting to apply a single compliance framework is highly complex because civilian markets trigger the EU AI Act and GDPR, while military use cases fall under national laws and NATO exemptions.	• **Implement Strategic Segmentation:** Clearly separate the product into two versions at the design phase (e.g., EPM-Civ and EPM-Mil) to ensure the civilian version gains CE compliance while the military version avoids unnecessary civilian bureaucracy.
• **"Black Box" Liability vs. Security:** Complying with the GDPR's requirement to explain the logic of an automated decision (as per the *Dun & Bradstreet* ruling) without revealing proprietary trade secrets or exposing vulnerabilities to adversaries.	• **Explainability by Design:** Instead of exposing the entire complex algorithm, design the system to output "key parameters" (e.g., "lens reflection detected") that satisfy transparency requirements without jeopardizing security or intellectual property.

References

1. Artificial Intelligence Act: https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=OJ:L_202401689
2. CJEU Case C-511/18 - La Quadrature du Net and Others, 2020. https://curia.europa.eu/juris/liste.jsf?language=en&num=C-511/18
3. The Cyber Resilience Act: https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=OJ:L_202402847
4. Digital Operational Resilience Act: https://eur-lex.europa.eu/eli/reg/2022/2554/oj/eng
5. General Data Protection Regulation: https://eur-lex.europa.eu/eli/reg/2016/679/oj/eng
6. NIS2 Directive: https://eur-lex.europa.eu/eli/dir/2022/2555/2022-12-27/eng
7. Law of the Republic of Lithuania on Cybersecurity https://www.e-tar.lt/portal/it/legalAct/5468a25089ef11e4a98a9f2247652cf4/asr
8. NATO AI Strategy, 2024. https://www.nato.int/en/about-us/official-texts-and-resources/official-texts/2024/07/10/summary-of-natos-revised-artificial-intelligence-ai-strategy
9. Treaty on European Union: https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:12012M/TXT
10. CJEU Case C-203/22 - Dun & Bradstreet Austria GmbH, 2025. https://curia.europa.eu/juris/liste.jsf?num=C-203/22
11. Data Strategy for the Alliance, 2025. https://www.nato.int/en/about-us/official-texts-and-resources/official-texts/2025/05/05/data-strategy-for-the-alliance
12. NATO's Data and Artificial Intelligence Review Board, 2022: https://www.nato.int/en/about-us/official-texts-and-resources/official-texts/2022/10/13/natos-data-and-artificial-intelligence-review-board
13. CJEU Case C-634/21 - SCHUFA Holding AG, 2023. https://curia.europa.eu/juris/liste.jsf?num=C-634/21
14. NATO Digital Backbone, 2024: https://www.nato.int/en/news-and-events/articles/news/2024/12/13/nato-digital-backbone
15. White Paper on Export Controls COM(2024) 25 final. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52024DC0025
16. Dual-Use Regulation (2021/821): https://eur-lex.europa.eu/eli/reg/2021/821/2024-11-08/eng
17. European Commission. AI Act Single Information Platform: https://ai-act-service-desk.ec.europa.eu/en
18. Action Plan on synergies between civil, defence and space industries. COM(2021) 70 final. https://www.eesc.europa.eu/en/our-work/opinions-information-reports/opinions/action-plan-synergies-between-civil-defence-and-space-industries

Editor
<<Main Editor>>

Contributors
Evaldas Bruze (MRU)
Giedrė Sabaliauskaitė (MRU)
R. Andrew Paskauskas (MRU)
Tomas Lavišius (MRU)
Saulius Kvedaravičius (MRU)

Reviewers
Evaldas Bružė (MRU)
Giedrė Sabaliauskaitė (MRU)
R. Andrew Paskauskas (MRU)
Tomas Lavišius (MRU)
Saulius Kvedaravičius (MRU)

Disclaimer