



AI-ASsisted cybersecurity platform empowering SMEs to defend against adversarial AI attacks

AIAS NEWSLETTER

Issue 7 | August 2026

AI systems find applications in various technical fields. However, their adoption exposes early users to vulnerabilities, such as data corruption, model theft, and adversarial samples. The lack of tactical and strategic capabilities to defend, identify, and respond to attacks on these AI-based systems is a significant concern. Adversaries exploit this vulnerability, creating a new attack surface that specifically targets Machine Learning and Deep Learning systems, posing a substantial threat to critical sectors like finance and healthcare. Addressing these challenges, the MSCA-funded AIAS project aims to conduct research on adversarial AI and develop an innovative security platform for organisations. This platform will employ adversarial AI defence methods, deception mechanisms, and explainable AI solutions to empower security teams, fortifying AI systems against potential attacks.

PROJECT COORDINATION

Prof. Christos Xenakis
School of Information and Communication
Technologies
Department of Digital Systems
University of Piraeus
Karaoli and Dimitriou 80,PC 18534, Piraeus,
Greece
Tel: +30 210 4142776
email: xenakis@unipi.gr

PROJECT DETAILS

Project number: 101131292
Project Website: aias-project.eu
Project start: 1st January 2024
Duration: 48 Months
Total cost: EUR 1564000
EC Contribution: EUR 1564000



This project has received funding from the European Union under HORIZON-TMA-MSCA-SE, Topic HORIZON-MSCA-2022-SE-01-01, Grant Agreement No 101131292.



AI-ASsisted cybersecurity platform empowering SMEs to defend against adversarial AI attacks

AIAS Objectives

Holistic Protection: Conceptualize and develop a service architecture integrating AI-empowered applications, deception mechanisms, and mitigation techniques towards the holistic protection of organizations against cyberattacks and adversarial AI.

Attack Scenarios: Design and develop a novel adversarial AI engine for creating attack scenarios tailored to the characteristics of the targeted organizations' hardware and software infrastructure.

Novel Intelligent Deception Methods: Design and implement novel intelligent deception methods based on high-interaction honeypots, digital twins, and virtual personas.

AI-based Methods for Protection: Design, develop, and assess AI-based methods for the detection and mitigation of cyberattacks including adversarial AI attacks as well as conceptualize and implement data collection and fusion methods.

XAI-based Recommendation Engine: Develop and verify explainable AI (XAI)-based recommendation engine empowering human-in-the-loop proactive decisions to thoroughly mitigate adversarial AI attacks.

Real-life Usage: Assess the functionality, effectiveness and efficiency of AIAS in real-life scenarios.



This project has received funding from the European Union under HORIZON-TMA-MSCA-SE, Topic HORIZON-MSCA-2022-SE-01-01, Grant Agreement No 101131292.



AI-ASsisted cybersecurity platform empowering SMEs to defend against adversarial AI attacks

AIAS Technical Highlights

During this period, the project moved on to develop tools to deceive and detect attacks against AI systems in action.

First, a survey and **taxonomy of AI attacks** was completed based on the existing literature across a wide range of application domains.

Specifically, rather than simply listing known attacks, the project partners classified them into a large-scale taxonomy based on distinctive characteristics, including the attacker's level of knowledge of the target AI/ML model, the stage at which the attack takes place, and the specific application domain involved.

Then, based on this understanding of the threats, the project also developed and implemented a set of **deception-based defence solutions** designed to protect AI systems in real-world operating conditions, along with a monitoring tool that collects and analyses data from both the target AI systems and the deception layer itself. Together, these form a comprehensive deception layer tailored to the scenarios defined earlier in the project.

Complementing the deception mechanisms, the project also provided a proof of concept for **security monitoring and analysis**, applying intelligent analytics techniques to transform raw system data into actionable information.



This project has received funding from the European Union under HORIZON-TMA-MSCA-SE, Topic HORIZON-MSCA-2022-SE-01-01, Grant Agreement No 101131292.



AI-ASsisted cybersecurity platform empowering SMEs to defend against adversarial AI attacks

News & Events

[Professor Christos Xenakis Participated in COcyber Panel Discussion on Europe's Cybersecurity Collaboration](#)

[AIAS MSCA Supports Cybersecurity Education through Gamified Learning](#)



[AIAS Participates in the Automation & Robotics Expo 2026](#)



This project has received funding from the European Union under HORIZON-TMA-MSCA-SE, Topic HORIZON-MSCA-2022-SE-01-01, Grant Agreement No 101131292.



AI-ASsisted cybersecurity platform empowering SMEs to defend against adversarial AI attacks

News & Events

AIAS Participates in the 2nd Hellenic University Hack – UniHack 2025

AIAS Participates in the “AI Enables Cybersecurity” Meetup



AIAS Participates in the 1st CyberHammer Hackathon 2026



This project has received funding from the European Union under HORIZON-TMA-MSCA-SE, Topic HORIZON-MSCA-2022-SE-01-01, Grant Agreement No 101131292.



AI-ASsisted cybersecurity platform empowering SMEs to defend against adversarial AI attacks

News & Events

AIAS Research and Expertise Featured at the 2nd AI in Financial Services Conference 2026



AIAS Research Presented at IEEE RTSI 2026



This project has received funding from the European Union under HORIZON-TMA-MSCA-SE, Topic HORIZON-MSCA-2022-SE-01-01, Grant Agreement No 101131292.



AI-ASsisted cybersecurity platform empowering SMEs to defend against adversarial AI attacks

AIAS Publication

Niculae, M., Suciu, G., Stanescu, V., Sachian, M. A., Farao, A., Sabazioti, A., ... & Karagiannis, S. (2024). Adversarial AI attack detection: a novel approach using explainable AI and deception mechanisms. In Smart Cities International Conference (SCIC) Proceedings (Vol. 12, pp. 623-647).

Suciu, G., Stalidi, C., & Popovici, E. C. (2025, June). Cross-Vertical Integration of AI, Blockchain, and IoT for Secure and Resilient Digital Health Ecosystems. In 2025 IEEE International Black Sea Conference on Communications and Networking (BlackSeaCom) (pp. 1-7). IEEE.

Stalidi, C., Popovici, E. C., & Suciu, G. (2025, August). Real-Time Digital Ecosystems: Integrating Virtual Personas and Digital Twins Through Microservices. In International Conference on Availability, Reliability and Security (pp. 128-141). Cham: Springer Nature Switzerland.

Ziras, G., Farao, A., Zarras, A., & Xenakis, C. (2025). From vulnerability to resilience: Adversarial training and real-time detection for AI security. Array, 100546.

Zarras, A., Kollarou, A., Farao, A., Bountakas, P., & Xenakis, C. (2025). Testing the limits: exploring adversarial techniques in AI models.



This project has received funding from the European Union under HORIZON-TMA-MSCA-SE, Topic HORIZON-MSCA-2022-SE-01-01, Grant Agreement No 101131292.



AI-ASsisted cybersecurity platform empowering SMEs to defend against adversarial AI attacks

AIAS Publication

- I. Politis, M. Bampatsikos, A. Zarras and C. Xenakis, "Trust Score Prediction for IoT Device Onboarding Using Transfer and Few-Shot Learning in Consumer Electronics," in IEEE Transactions on Consumer Electronics
- Kotsiopoulos, T., Radoglou-Grammatikis, P., Lekka, Z. et al. Defending industrial internet of things against Modbus/TCP threats: A combined AI-based detection and SDN-based mitigation solution. Int. J. Inf. Secur. 24, 157 (2025). <https://doi.org/10.1007/s10207-025-01076-2>
- Farao, A., Bolgouras, V., Zarras, A., & Xenakis, C. Cybersecurity Challenges and Pitfalls in 6G Networks.
- Petihakis, G., Farao, A., Bountakas, P., Sabazioti, A., Polley, J. and Xenakis, C., 2024, July. AIAS: AI-ASsisted cybersecurity platform to defend against adversarial AI attacks. In Proceedings of the 19th International Conference on Availability, Reliability and Security (pp. 1-7).
- Bampatsikos M, Politis I, Ioannidis T, Xenakis C. Trust Score Prediction and Management in IoT Ecosystems Using Markov Chains and MADM Techniques. IEEE Transactions on Consumer Electronics. 2025 Jan 17.



This project has received funding from the European Union under HORIZON-TMA-MSCA-SE, Topic HORIZON-MSCA-2022-SE-01-01, Grant Agreement No 101131292.



AI-ASsisted cybersecurity platform empowering SMEs to defend against adversarial AI attacks

AIAS Publication

Lacalle I, Cuñat S, Farao A, Xenakis C, Xenakis D, Palau CE. Deception Mechanisms for Cyber-Security Enhancement in the Internet of Things. In 2024 IEEE 29th International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD) 2024 Oct 21 (pp. 1-7). IEEE.

Sotiropoulos, A., Xenakis, D., Kalpaktoglou, G., Vittorakis, Y., & Dolapsaki, M. (2026, January 13). On the resource consumption of XR-enabled Digital Twin Applications in the edge-cloud continuum.

Upcoming Technical Deliverables

D4.1-AI-based Detection of Adversarial AI Attacks (December/2026)

D4.2-Mitigation of Adversarial AI Attacks & XAI (June/2027)

D5.1-Platform Integration (June/2027)

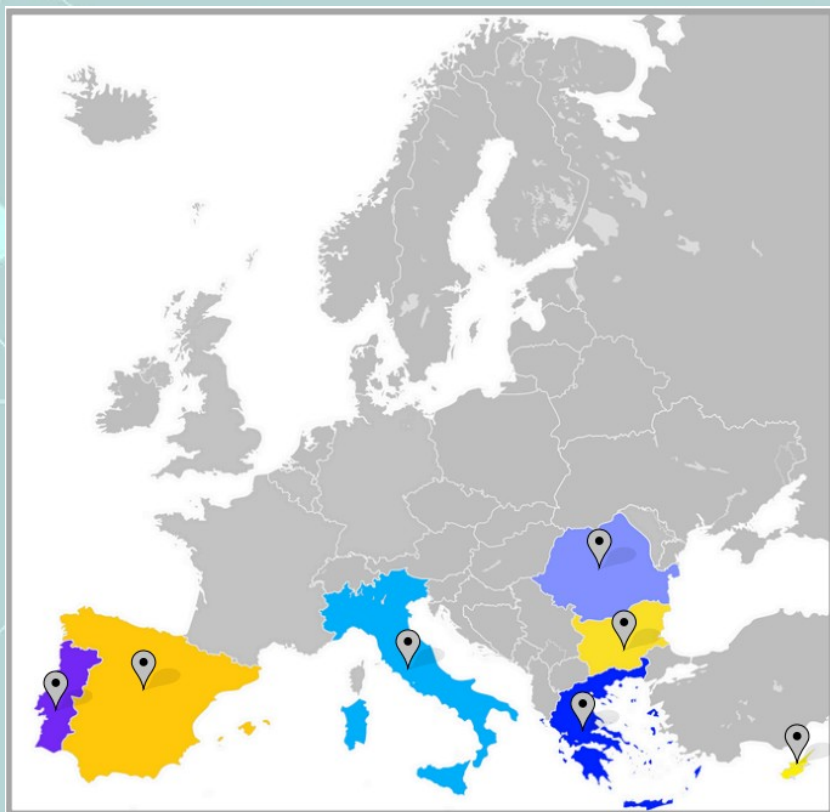


This project has received funding from the European Union under HORIZON-TMA-MSCA-SE, Topic HORIZON-MSCA-2022-SE-01-01, Grant Agreement No 101131292.



AI-ASsisted cybersecurity platform empowering SMEs to defend against adversarial AI attacks

Meet the Consortium



Follow us for our latest news!

www.aias-project.eu

[@AIAS.MSCA](https://www.facebook.com/AIAS.MSCA)

[@AIAS MSCA](https://www.linkedin.com/company/aias-msca)

[@AIAS MSCA](https://www.x.com/AIAS_MSCA)



This project has received funding from the European Union under HORIZON-TMA-MSCA-SE, Topic HORIZON-MSCA-2022-SE-01-01, Grant Agreement No 101131292.