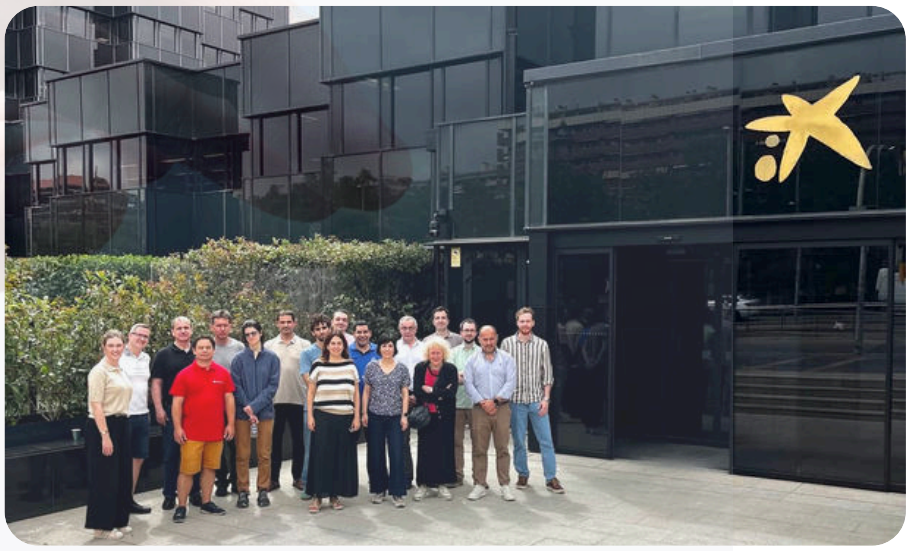
AI4CYBER NEWSLETTER

Edition 1, N3

August 2025

AI4CYBER FINAL YEAR

Launched in September 2022, AI4CYBER is an EU-funded research project under Horizon Europe **(GA ID: 101070450)** which brings together 13 partners from 7 EU Member States to develop and test smart cybersecurity services averaging AI and Big Data. The AI4CYBER ecosystem framework brings together 11 AI-based services of different cybersecurity areas. AI4CYBER aims to enhance the security and resilience of critical systems against advanced and AI-powered cyberattacks. It plans to do so by advancing the existing state-of-the-art AI techniques to create specialised solutions for prevention, detection, response and threat intelligence & integrate them in one framework that will be modular and complemented with Trustworthy AI services.



Final Consortium Meeting in Barcelona, June 12-13, 2025

CONTENT

Funded by the European Union

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101070450. Disclaimer: 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 European Commission can be held responsible for them.

AI4CYBER

Congratulations from the coordinator

AI4CYBER has been one of the EU-funded projects pioneer in the experimentation with the use of AI and generative AI for enhancing different cybersecurity services. Our research included multiple areas, including code vulnerability scanning and fixing (AI4VULN and AI4FIX), Cyber Threat Intelligence analysis and augmentation (AI4CTI), advanced attack simulation (AI4SIM), advanced threat detection (AI4FIDS), alert triage (AI4TRIAGE), security orchestration automation and response (AI4SOAR), response adaptability (AI4ADAPT), honeynets and deception solutions (AI4DECEIVE), and, last but not least, anonymous sharing of incident information (AI4COLLAB).

Erkuden Rios Velasco, PhD.
Project Manager, TECNALIA

After 3 years of successful collaboration in the project research work, key advances in the use of AI and Gen AI for cybersecurity and resilience capabilities of the systems were achieved. The research results in all of our components have demonstrated that the AI and the generative AI (in particular Large Language Models - LLMs) enable smarter, faster, and more efficient security against advanced attacks and offensive AI. Furthermore, following the requirements by the EU AI Act, the project has also investigated in parallel how our components could be ensured to be trustworthy, with a focus on explainability, fairness, and security aspects, all together under the umbrella of the so-called TRUST4AI component, which could also be use by other AI system developers in the task of ensuring trustworthy AI.

AI4CYBER has delivered the AI4CYBER framework, a collection of the eleven AI-powered cybersecurity tools named above, which can be adopted individually, in packages or as a whole. In fact, 3 of these tools are also sets of tools that can be adopted individually: AI4SIM, a collection of multiple types of attack simulators, including AML simulators; AI4FIDS, a set of advanced federated IDS, applicable to diverse detection problems; and finally, TRUST4AI that includes 3 solutions, TRUST4AI.XAI, TRUST4AI.Fairness and TRUST4AI.Security.

Funded by the European Union

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101070450. Disclaimer: 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 European Commission can be held responsible for them.

AI4CYBER

AI4CYBER SOLUTIONS (PART 1)

- **AI4VULN** is a static source code analyzer tool that can detect vulnerability problems in the source code. It uses symbolic execution, which means that it simulates the execution of the program without real execution and this way, it can detect runtime vulnerability issues in the code.

- **AI4FIX** enables to shift the fixing of the vulnerability much earlier in the software development flow, which in turn saves development time and reworks. The tool can be used with online and offline Large Language Models (LLM) and it is able to generate tests to improve code coverage.

- **AI4CTI** is a LLM-based Cyber Threat Intelligence (CTI) knowledge analyzer and aggregator, which primary goal is to gather insights from public CTI sources and to identify the techniques and attack paths commonly used by threat actors in sophisticated cyber-attacks.

- **AI4SIM** is a modular framework designed to simulate a range of AI-powered attacks, adversarial attacks, and advanced complex attacks. It enables security researchers, developers, and system operators to rigorously test the resilience of systems against next-generation threats that leverage AI. The tool is structured around eight complementary modules, each focusing on specific types of AI-driven offensive techniques:
  - Manager of Attack Generators and dashboard
  - GAN-based Network Fuzzer
  - AI-based KNX Fuzzer
  - A specialized AI-driven fuzzer for KNX/IP protocols
  - MAIP module dedicated to crafting adversarial attacks against Intrusion Detection Systems (IDSs).
  - Adversarial Attack Generator designed to create adversarial datasets and traffic against AI-based intrusion detection systems.
  - LLM-based Attack generator that leverages Large Language Models (LLMs) to automate complex attack strategies.
  - Advanced Attacks with MITRE Caldera, enabling automated execution of multi-stage attacks.
  - AI-based Penetration Testing using DeepExploit.

- **AI4FIDS** is a high-performance solution designed to detect a broad spectrum of attacks across network, host, and log data. Its core feature is federated learning, allowing multiple clients to collaboratively train an intrusion detection model without sharing their raw data.

Funded by the European Union

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101070450. Disclaimer: 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 European Commission can be held responsible for them.

AI4CYBER

AI4CYBER SOLUTIONS (PART 2)

- **AI4TRIAGE** is an AI-based root cause analysis and alert triage solution to prioritize events and best focus the threat mitigations and response.

- **AI4SOAR** is an AI-powered Security Orchestration, Automation and Response solution capable to deploy multiple security controls at different layers of the system for better react against cyber incidents and attacks.

- **AI4DAPT** is an "intelligent" Reinforcement Learning (RL) solution with the main objective of optimizing the incident response adaptation. The component is designed to defend the victim system against certain types of attacks (advanced attacks in the AI4CYBER nomenclature) which are composed of a chain of different atomic attacks.

- **AI4COLLAB** represents a transformative step forward in Cyber Threat Intelligence (CTI) sharing, seamlessly integrating MISP and OpenCTI platforms to enable a secure, real-time, and scalable exchange of threat data. The platform ensures that CTI flows efficiently between stakeholders while complying with GDPR and preserving data utility.

- **AI4DECEIVE** is the solution for deception and it includes an Intelligent Deceive Decision Engine (IDDE) that applies game theory to determine the optimal type and number of honeypot instances needed to mitigate detected threats against a system.

- **TRUST4AI.Security** is a solution to facilitate the search for all possible adversarial threats and protections relevant to the AI model under study. Therefore, it significantly facilitates the task of risks assessment in AI-based systems in regard to AI robustness and security.

- **TRUST4AI.FAIRNESS** proposes fairness evaluation and mitigation mechanisms to ensure that AI-based systems do not introduce bias and discrimination in the process on AI analyses.

- **TRUST4AI.XAI** provides a user-friendly dashboard to visualize and interpret AI decisions, integrates easily with external models, and introduces new explainability metrics to evaluate feature importance even before training. Designed for cybersecurity, it helps analysts understand why threats are detected, ensuring AI systems are more reliable and robust in real-world scenarios.

Funded by the European Union

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101070450. Disclaimer: 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 European Commission can be held responsible for them.

AI4CYBER

MAIN AI4CYBER ACTIVITIES IN FINAL YEAR

As we mark the final year of the AI4CYBER project, we are proud to highlight significant achievements in developing next-generation AI-based services for cybersecurity. Over the past year, we've made substantial progress in the development of the components of the ecosystem framework that enhances the robustness and resilience of critical systems against advanced cyber threats. AI4CYBER's partners were active in various events in the past months, including major conferences, worskhosp and summer schools. These engagements ensured strong visibility for the project while fostering new collaborations and expanding our cybersecurity academic and professional community.

AI4CYBER's partners attended various events in the past months to present and discuss the project outcomes, including:

- Organisation of STAM Workshop 2025, August 2024 and August 2025.
- AI4CYBER lectures in TAROT Summer School, June 30 -July 2nd 2025
- Participation at the Basque CyberIndustry Congress, June 17th 2025
- Participation at Barcelona Cybersecurity Congress (BCC), May 2024 and May 2025
- Organisation and hosting of the 3rd Workshop of the European Cluster for Securing Critical Infrastructures, April 29-30th 2025
- Booth at InCyber Forum, Booth, April 1-3rd 2025
- Co-organisation with NCP and hosting of the of the "Jornada Cluster-3 Civil Security, Horizon Europe call-2025. Encuentro actores vascos de la innovación en seguridad", July 1st 2025
- Participation at the EU-CIP Annual Conference, November 13th 2024
- Booth at European Cyber Week Rennes, November 18-21st 2024
- Participation at the Design, Verification and Validation of IoT Systems Workshop, November 6th 2024
- Presentation at 88th Thessaloniki International Fair 2024, September 8th 2024
- Organisation of the IEEE CSR Workshop on Electrical Power and Energy Systems Security, Privacy and Resilience (EPES-SPR), September 4th 2024
- Presentation at 20th International Conference on Artificial Intelligence Applications and Innovations, June 28th 2024
- Keynote at Madeira Digital Transformation Week, June 26th 2024

Funded by the European Union

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101070450. Disclaimer: 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 European Commission can be held responsible for them.

AI4CYBER

READ THE AI4CYBER BLOGPOSTS

Most recent Blogposts :

Security of Artificial Intelligence by AI4CYBER, Aug 2025

From Trusted AI to Resilient Critical Systems, Aug 2025

Ensuring Fairness in Large Language Models: An Emerging Critical Dialogue, Aug 2025

Ensuring Business Continuity in Healthcare : Managing Cyber-Physical Threats, Jul 2025

Aligning AI Regulation with Cybersecurity for a Competitive Europe, Jun 2025

TRUST4AI.xAI: Enhancing AI Transparency and Trustworthiness in Cybersecurity, May 2025

StatAvg: An Open-source Technique for Mitigating Client Heterogeneity in Federated Learning-based Intrusion Detection Systems, Mar 2025

Enhancing Automated Cybersecurity Incident Response with AI4SOAR, Feb 2025

AI4CYBER CONSORTIUM

tecnalia MEMBER OF BASQUE RESEARCH & TECHNOLOGY ALLIANCE

FRONTENDART

ΠΑΝΕΠΙΣΤΗΜΙΟ ΔΥΤΙΚΗΣ ΜΑΚΕΔΟΝΙΑΣ

THALES Building a future we can all trust

CaixaBank

EOS European Organisation for Security

iTTi

ΔΕΗ

SEARCH-LAB Security Evaluation Analysis and Research Laboratory

UNIDADE LOCAL DE SAÚDE ALENTEJO CENTRAL

MINDS

montimage

PDM

Contact

erkuden.rios@tecnalia.com

@Ai4Cyber

@AI4CYBER

@AI4CYBER Project

https://ai4cyber.eu

Funded by the European Union

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101070450. Disclaimer: 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 European Commission can be held responsible for them.