



AI-BASED CCAM: TRUSTWORTHY, EXPLAINABLE, AND ACCOUNTABLE

The Autonomous Main Event 2024

Spotlight Session – Driving the Future: The Rise of Software-Defined Vehicles

Dr. Georg Stettinger

24.09.2024

Motivation

Connected and Cooperative Automated Mobility (CCAM) is blooming thanks to **Artificial Intelligence (AI)**

- CCAM solutions have benefited from the applicability of AI-based perception, situational awareness, and decision-making components
- learning highly complex transformations that operate over input sensor data and produces end-commands (steering angle, throttle).

... unveiled the fact

- **Black Boxes** for their critical lack of transparency and interpretability.
- **AI can be unfair** and biased, expose private data, and be extremely sensitive to unexpected inputs

Trustworthy AI is the next mandatory step of technology development

- trustworthy AI requires exploring trade-offs among other equally important properties: **robustness, privacy, explainability, accountability, and ethics**

AITHENA PROJECT

Call: Safe, Resilient Transport and Smart Mobility services for passengers and goods

Topic: HORIZON-CL5 2022 D6-01-05: Artificial Intelligence (AI): Explainable and trustworthy concepts, techniques and models for CCAM

Type of Action: Research and Innovation Action

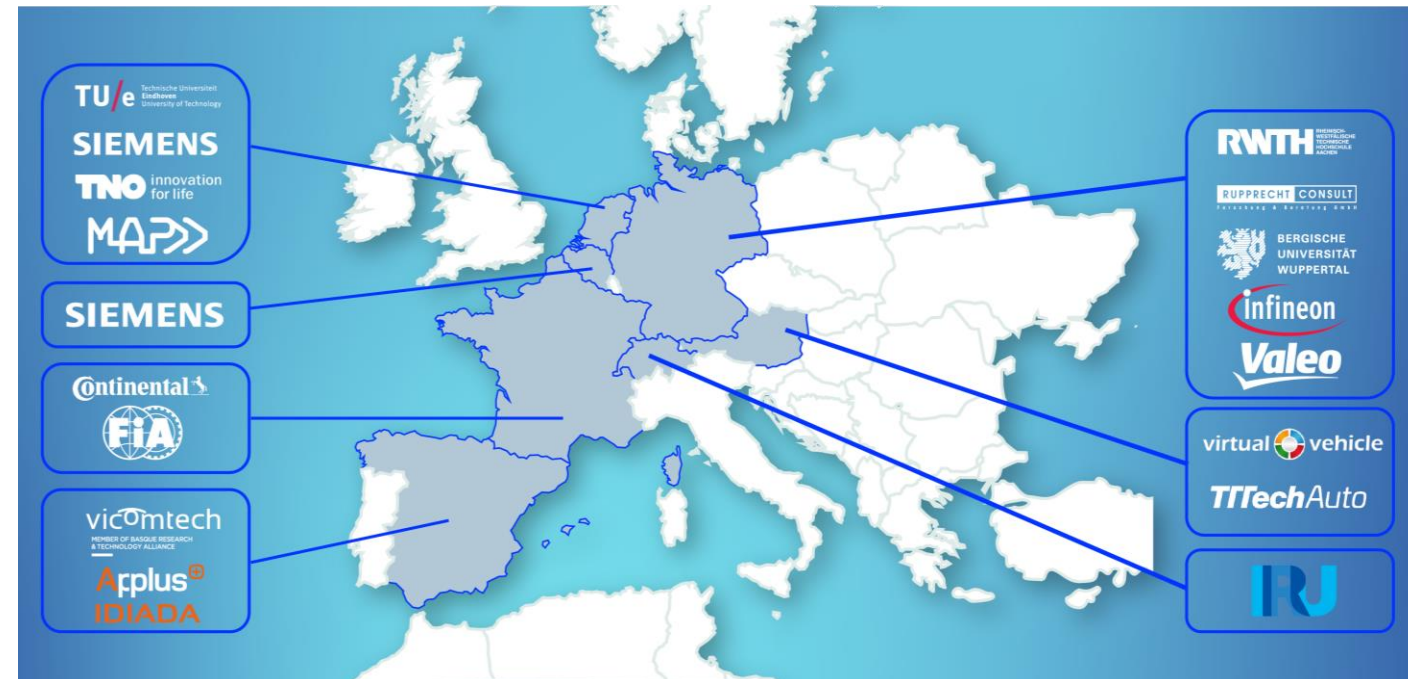
Coordinator: VICOMTECH (Spain)

Consortium: 17 partners

Start: 1 November 2022

Duration: 36 months

Budget: 5.999.549 €

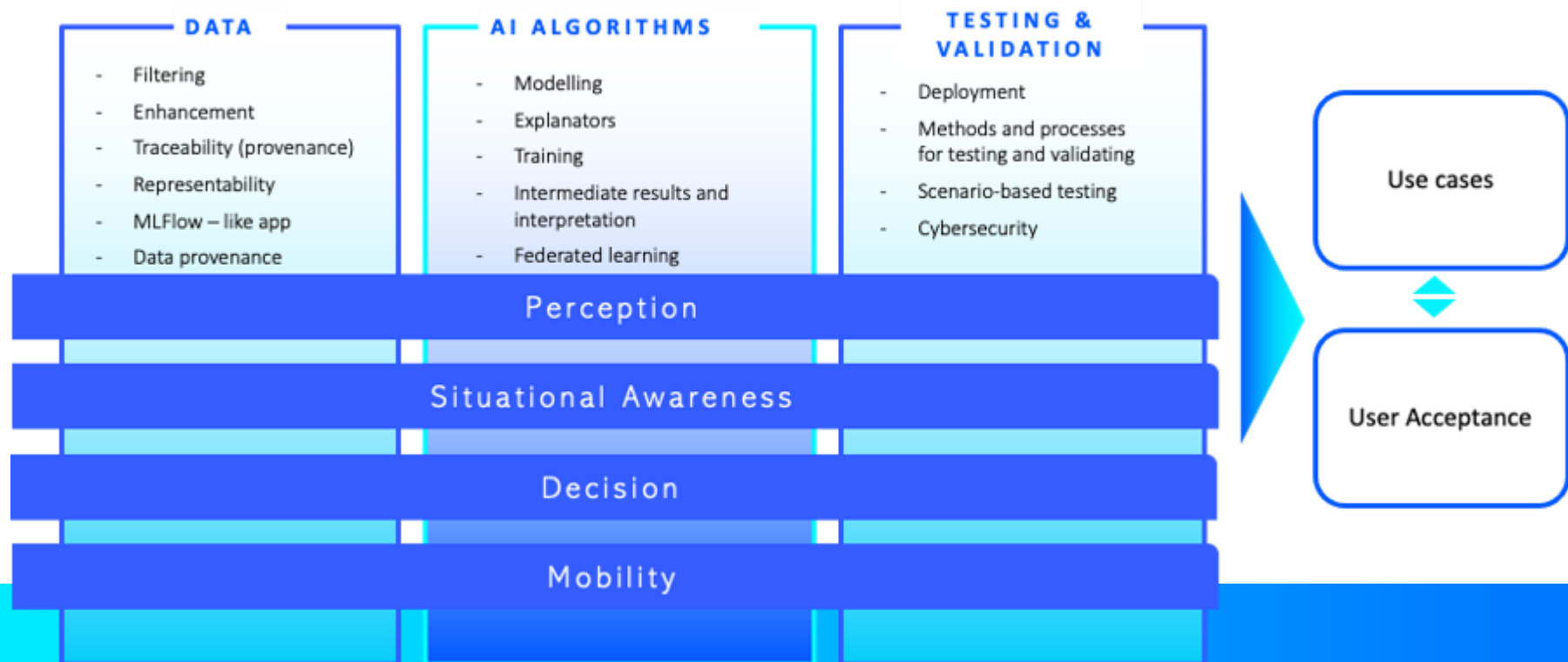


APPROACH

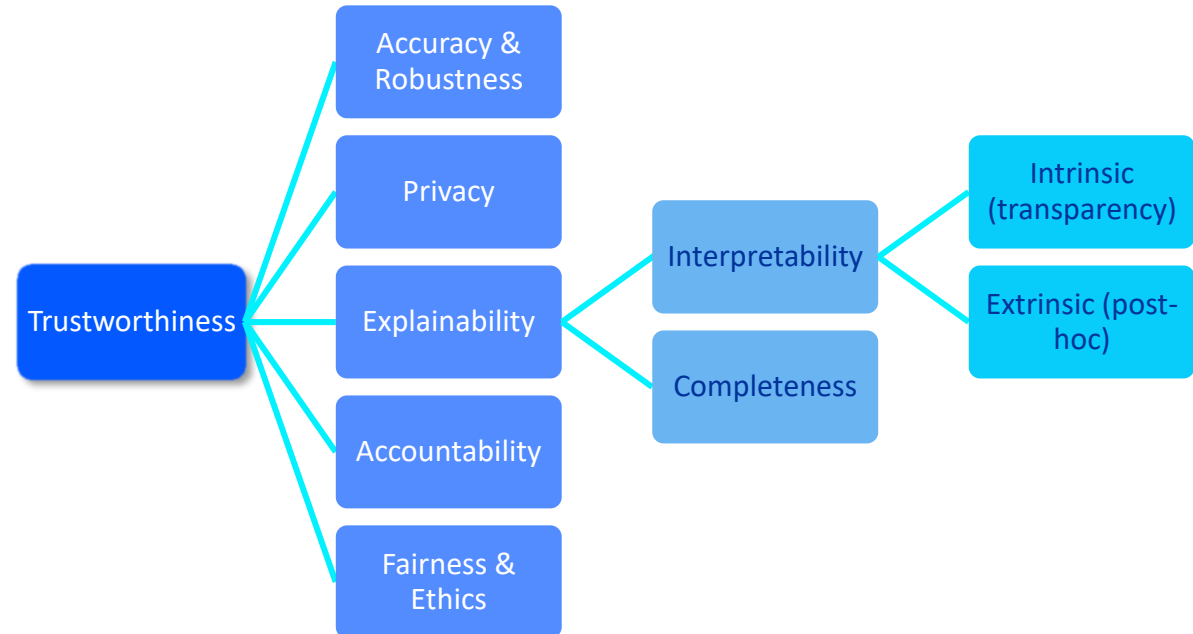
AITHENA will face **methodological** and development challenges for the **creation and integration** of **XAI-based models** and systems into **CCAM applications** (use cases: perception, situational awareness, decision-making and traffic management).

AITHENA will provide a human-centered METHODOLOGY towards the evolution of the three main AI pillars: DATA, AI MODELS and TESTING

Human centric Approach & Ethics



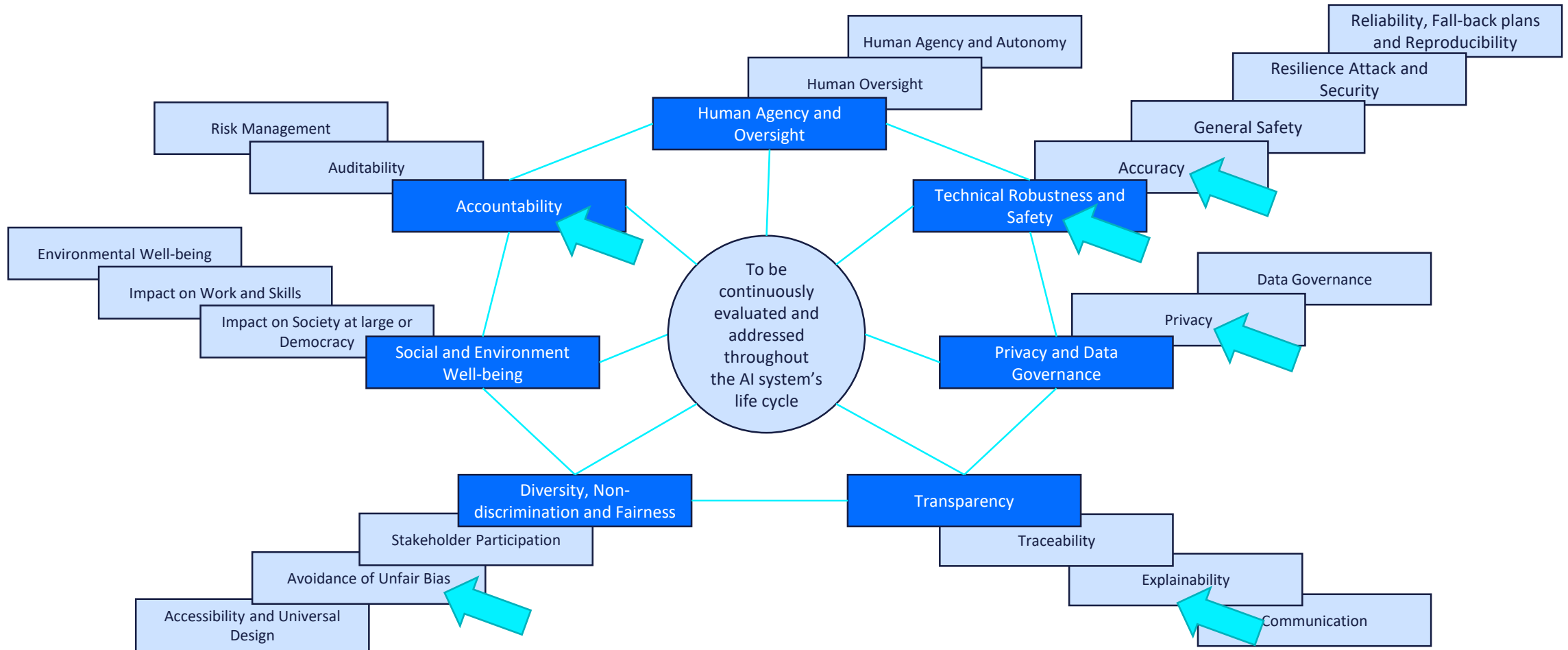
GENERAL REQUIREMENTS – AREAS OF ACTIVITY



Requirements of Trustworthy AI – Link to the AI Act

7 requirements tightly interconnected

Athena focus topics



VALUE PROPOSITION

Harmonized Methodology



CCAM

Human Centric Approach

DATA

- Real & Synthetic Data generation
- Ontology, Semantics & metadata
- Representability
- Life Cycle and provenance
- Devs Op

X-AI

- physics-informed neural networks
- Deep hybrid learning
- Situational Awareness
- Reinforcement Learning
- Explanators

V2V

- KPI Definition
- Scenario based testing
- Extension of current V&V Methods

Standardisation (EuroNCAP), Testing & Validation



Open Tools

Open DDBB



Perception



Understanding



Decision



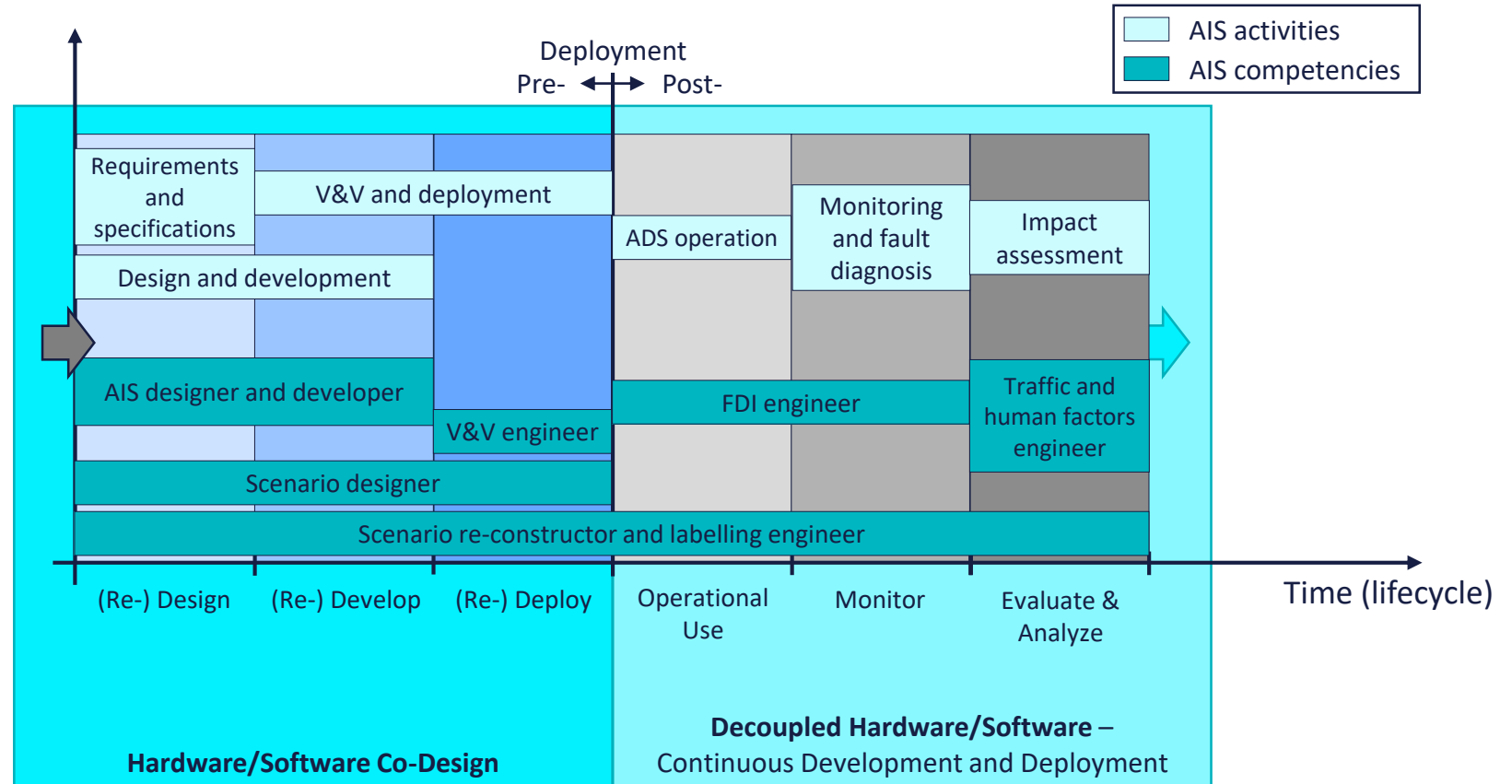
Traffic

Harmonized Methodology

AI lifecycle acts as a game changer

HW/SW co-design vs. decoupling HW/SW

- The AIS life cycle incorporates the **entire AIS value chain**
- Software-defined-Vehicles (SdV) and AI in particular **changed the way of working completely**
 - New focus element: post-deployment phase
 - Continuous development and deployment (software)
- Pre-deployment phase is driven by **hardware/software co-design** principles while the post-deployment phase is **hardware/software decoupled**



SPECIFIC OBJECTIVES

- Explainable, privacy-preserving and traceable data management Tools
- Development Framework (“DevOps” like) to ensure data and model lifecycle tracking and management
- Digital Twin for data generation

- Methodology and good practices for AI based CCAM solution development and testing
- XAI models – research and development of explainable AI models
- Development of human-centric AI solutions for future CCAM applications strengthening user acceptance, explainability and trustworthiness

- Testing and validation procedures – methodology for extending HEADSTART methodology to include AI based functions and systems
- Trusted AI Key Performance Indicators for CCAM components

CCAM

- Trustworthy and Robust Perception systems
- Human Understandable Situation Awareness System including driver state
- Explainable Driving Decision methods
- AI-based traffic analysis module

References

[1] G. Stettinger, P. Weissensteiner and S. Khastgir, "Trustworthiness Assurance Assessment for High-Risk AI-Based Systems," in IEEE Access, vol. 12, pp. 22718-22745, 2024, doi: 10.1109/ACCESS.2024.3364387.



THANK YOU!

Website: <https://aithena.eu/>

X (Twitter): <https://twitter.com/Althena>

LinkedIn: <https://www.linkedin.com/company/aithena-eu-project/>



Funded by
the European Union

Project funded by



Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Swiss Confederation

Federal Department of Economic Affairs,
Education and Research EAER
State Secretariat for Education,
Research and Innovation SERI