NTNU

Norwegian University of Science and Technology

# Sustainable Real-time Anomaly Detection for Practical Applications

DC12

**Athanasia Kollarou**

Norwegian University of Science and Technology

TUAI Summer School · Gliwice · July 2026

HORIZON-MSCA-2023-DN-01 · GA No 101168344 · WP5







# Objectives and Expected Results

*Real-time anomaly detection provides immediate insight into data streams in cybersecurity, industrial monitoring, healthcare, finance and smart infrastructure. The challenge: achieve its goals while assuring resource efficiency, scalability, privacy, resilience, transparency, interpretability and cross-domain applicability.*

**Sustainability dimensions (GA)**

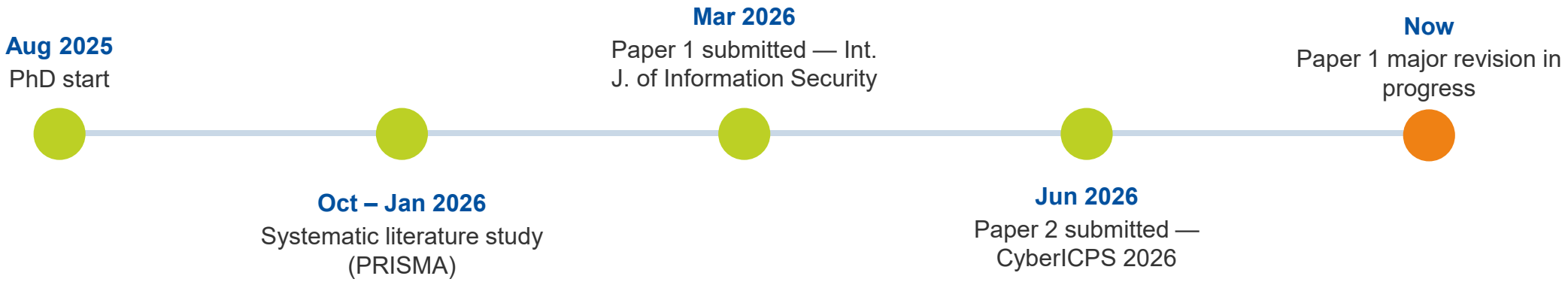
Resource efficiency · Scalability · Data privacy · Resilience · Transparency · Interpretability · Cross-domain applicability

**EXPECTED RESULTS (i–v)**

- **i** Comprehensive analysis of state-of-the-art real-time anomaly detection, focused on sustainability aspects
- **ii** Development of a sustainable real-time anomaly detection solution
- **iii** Demonstration on safety-critical cyber-physical systems
- **iv** Extensive evaluation using a set of sustainability metrics
- **v** Dissemination through research papers and presentations

PROGRESS UP TO DATE

# Project Submissions




**Status:** two publications produced — a deployability survey (under major revision) and a security-requirements paper (submitted).

PUBLICATION 1 · EXPECTED RESULT (I)

# A Survey and a Deployability Assessment

**Towards Deployable Anomaly Detection in Industrial CPS/ICS: A Survey and Deployability Assessment**

Kollarou et al. · International Journal of Information Security · under major revision

## Motivation

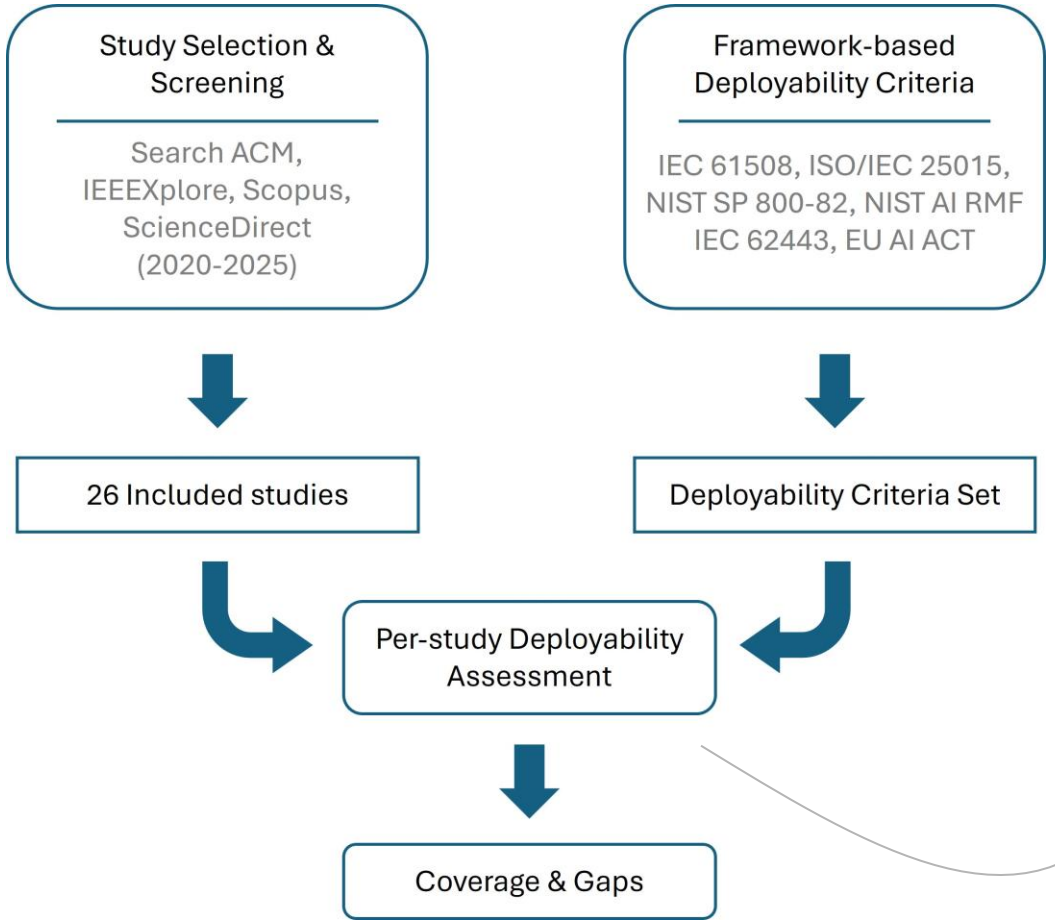
- How are anomaly detection solutions approaching deployability in CPS/ICS?
- What are properties that support deployability claims?.

## Contributions

- A set of nine **deployability criteria** derived from safety, security and AI-governance frameworks.
- Evidence-based assessment and review of **26 implementation studies** (fully / partially / not considered).
- A gap map between reported performance and deployment readiness.

PUBLICATION 1

# Methodology: PRISMA Review & Criteria Derivation




| Research Work | Family | ACC | OR | AD | RTR | RE | AR | CONF | EXP | VER |
|---|---|---|---|---|---|---|---|---|---|---|
| Zhu et al. [57] | Mathematical | ● | ● | ● | ● | ◐ | ● | ○ | ◐ | ◐ |
| Zhang et al. [55] | Mathematical | ● | ● | ◐ | ● | ◐ | ○ | ○ | ◐ | ◐ |
| Jin et al. [24] | Mathematical | ● | ● | ◐ | ● | ◐ | ○ | ○ | ◐ | ◐ |
| Hosseinzadehtaher et al. [17] | Mathematical | ● | ● | ◐ | ● | ○ | ● | ○ | ◐ | ○ |
| Birihanu and Lendák [6] | Machine Learning | ● | ◐ | ◐ | ◐ | ◐ | ○ | ○ | ● | ◐ |
| Yang et al. [51] | Machine Learning | ● | ◐ | ◐ | ● | ○ | ● | ○ | ○ | ○ |
| Mokhtari et al. [32] | Machine Learning | ● | ○ | ○ | ○ | ○ | ◐ | ○ | ○ | ◐ |
| Liu et al. [29] | Deep Learning | ● | ◐ | ◐ | ○ | ○ | ● | ○ | ○ | ◐ |
| Yang et al. [52] | Deep Learning | ● | ● | ● | ● | ● | ◐ | ○ | ○ | ◐ |
| Zare et al. [54] | Deep Learning | ● | ◐ | ◐ | ● | ◐ | ● | ○ | ○ | ○ |
| Liao et al. [28] | Deep Learning | ● | ◐ | ◐ | ◐ | ◐ | ◐ | ○ | ◐ | ◐ |
| Aslam et al. [5] | Deep Learning | ● | ◐ | ◐ | ◐ | ◐ | ◐ | ○ | ◐ | ◐ |
| Tang et al. [48] | Deep Learning | ● | ◐ | ○ | ○ | ◐ | ◐ | ○ | ● | ○ |
| Macas et al. [30] | Deep Learning | ● | ◐ | ○ | ○ | ◐ | ◐ | ○ | ○ | ◐ |
| Kea et al. [25] | Deep Learning | ● | ◐ | ◐ | ● | ● | ◐ | ● | ○ | ◐ |
| Sahin et al. [43] | Deep Learning | ● | ◐ | ○ | ○ | ◐ | ◐ | ○ | ○ | ○ |
| Do Thu Ha et al. [12] | Deep Learning | ● | ○ | ◐ | ◐ | ◐ | ◐ | ○ | ● | ◐ |
| Zhou et al. [56] | Deep Learning | ● | ● | ◐ | ● | ◐ | ◐ | ○ | ○ | ◐ |
| Huong et al. [19] | Deep Learning | ● | ● | ● | ● | ● | ◐ | ● | ○ | ○ |
| Jeffrey et al. [23] | Hybrid | ● | ○ | ◐ | ◐ | ◐ | ○ | ○ | ○ | ○ |
| Hu et al. [18] | Hybrid | ● | ● | ◐ | ◐ | ◐ | ● | ○ | ◐ | ◐ |
| Hao et al. [13] | Hybrid | ● | ◐ | ◐ | ● | ● | ○ | ○ | ○ | ○ |
| Asefi et al. [4] | Hybrid | ● | ● | ◐ | ◐ | ◐ | ● | ○ | ◐ | ◐ |
| Wu et al. [50] | Hybrid | ● | ◐ | ○ | ○ | ◐ | ◐ | ○ | ● | ◐ |
| Raman and Mathur [41] | Hybrid | ● | ◐ | ○ | ● | ◐ | ● | ○ | ◐ | ◐ |
| Feng and Tian [9] | Hybrid | ● | ● | ○ | ◐ | ◐ | ○ | ○ | ○ | ◐ |
| **Fully (●)** | | 26 | 10 | 3 | 12 | 4 | 8 | 2 | 4 | 0 |
| **Partially (◐)** | | 0 | 13 | 16 | 8 | 18 | 12 | 0 | 9 | 18 |
| **Not (○)** | | 0 | 3 | 7 | 6 | 4 | 6 | 24 | 13 | 8 |

● Fully considered: explicitly targeted as a design goal and supported by design choices, or evaluated with appropriate metrics.
◐ Partially considered: mentioned or implied through a design choice but not explicitly targeted or evaluated.
○ Not considered: no discussion, design choice, or evaluation related to the criterion.

PUBLICATION 1

# Nine Deployability Criteria

**ACC**
**Detection Accuracy**

**OR**
**Operational Resilience**

**AD**
**Adaptability**

**RTR**
**Real-Time Responsiveness**

**RE**
**Resource Efficiency**

**AR**
**Adversarial Robustness**

**CONF**
**Data Confidentiality**

**EXP**
**Explainability**

**VER**
**Verification**

*These criteria operationalise the GA's sustainability dimensions — the design and evaluation frame for the solution phase.*

PUBLICATION 1

# Key Contributions

### Detection Accuracy ✓

All 26 studies consider detection accuracy, but coverage of the remaining criteria varies widely.

### Criteria co-occur

Clear co-occurrence patterns emerge between criteria addressed simultaneously, for example, real-time responsiveness paired with resource efficiency.

### Deployability criteria

The nine criteria guide detector development, and extend naturally into an evaluation framework for assessing the deployability of existing detectors.

### Adversarial robustness & data confidentiality are among the least supported criteria, motivating Publication 2…

PUBLICATION 2

# Securing the Detectors Themselves

## From Adversarial Threat Models to Security Requirements for Anomaly Detectors

Kollarou et al. · submitted to CyberICPS 2026 · under review

## Motivation

- ML-based **anomaly detectors are themselves attack targets**: poisoning, evasion and privacy attacks can hide malicious behavior or leak sensitive data.
- Adversarial-ML literature rarely addresses the specific characteristics of anomaly detection from a **security-requirements perspective**.

## Contributions

- Structured **threat analysis** along attacker **objective**, **knowledge**, **capabilities** and **timing**.
- **Lifecycle mapping** of attacks & defenses, qualitative risk evaluation, three deployment use cases.
- A set of **security requirements** for adversarially robust detectors, each linked to attacks and defenses.

PUBLICATION 2

# Securing the Detectors Themselves

## From Adversarial Threat Models to Security Requirements for Anomaly Detectors

Kollarou et al. · submitted to CyberICPS 2026 · under review

| | | | | |
|---|---|---|---|---|
| Poisoning Attacks | Data Poisoning Attacks | Availability data poisoning | Targeted data poisoning | Backdoor data poisoning |
| | Model Poisoning Attacks | Availability model poisoning | Targeted model poisoning | Backdoor model poisoning |
| Evasion Attacks | White-box Attacks | Optimization-based methods | Universal evasion attacks | Physically realizable attacks |
| | Black-box Attacks | Score-based attacks | Decision-based attacks | |
| Privacy Attacks | Data Privacy Attacks | Data reconstruction | Membership inference | Property inference |
| | Model Privacy Attacks | Model extraction | | |

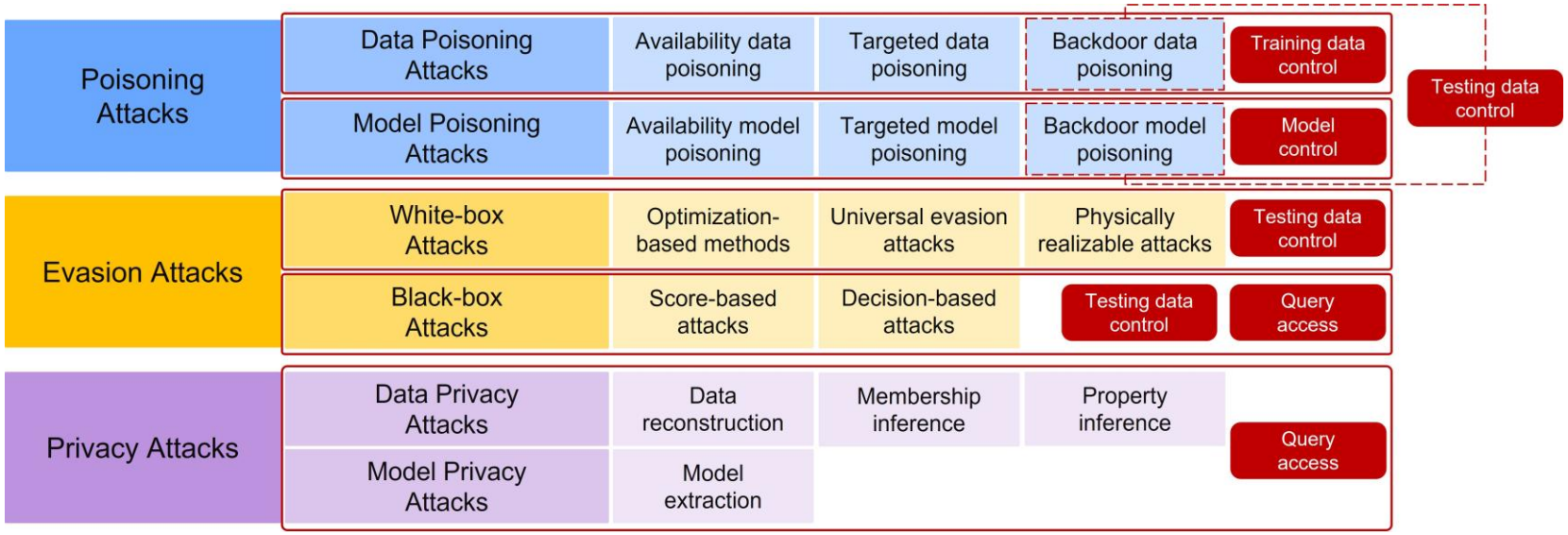
*Taxonomy: NIST AI 100-2e 2025*

PUBLICATION 2

# Securing the Detectors Themselves

## From Adversarial Threat Models to Security Requirements for Anomaly Detectors

Kollarou et al. · submitted to CyberICPS 2026 · under review



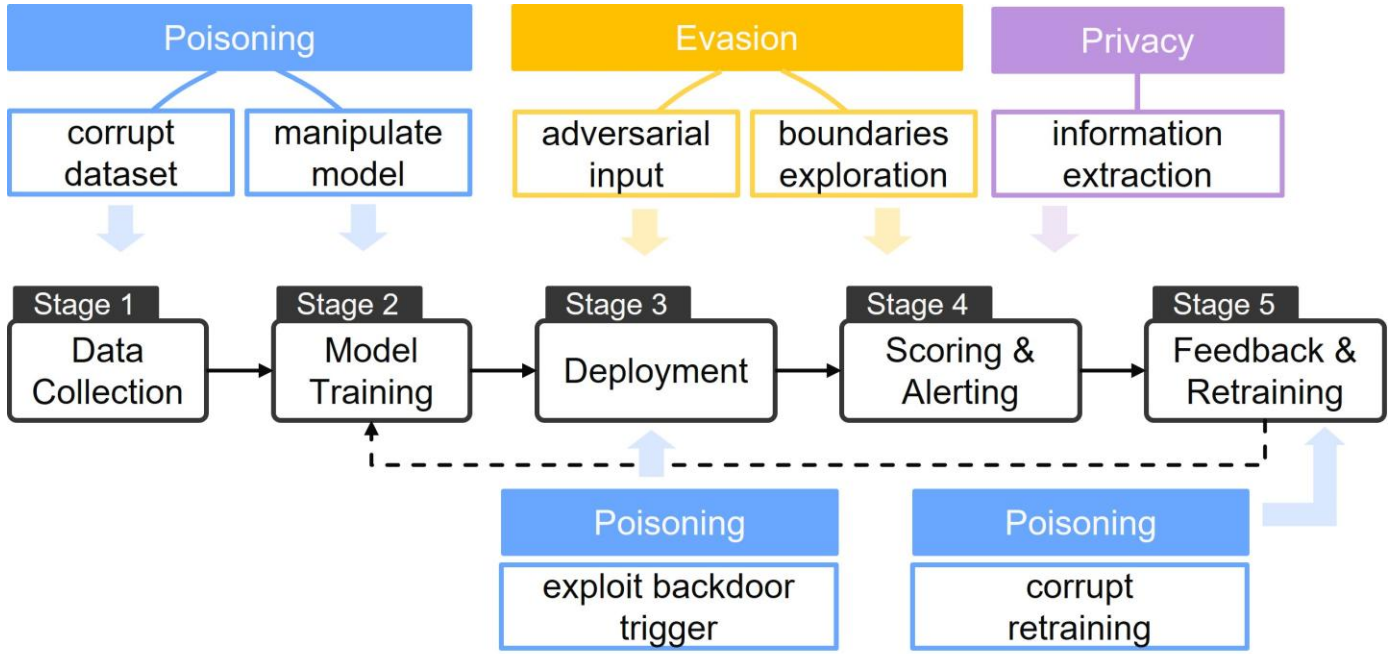

*Taxonomy: NIST AI 100-2e 2025*

PUBLICATION 2

# Securing the Detectors Themselves

**From Adversarial Threat Models to Security Requirements for Anomaly Detectors**

Kollarou et al. · submitted to CyberICPS 2026 · under review



PUBLICATION 2

# Approach: STORE-based Requirements Engineering

**Threat analysis** 15 attack subclasses in 3 families — poisoning, evasion, privacy — mapped to the AD lifecycle stage where they occur

**Defense review** Defense strategies organised by lifecycle stage; no single mechanism secures the detector end-to-end

**Risk evaluation** Qualitative probability × impact rating per attack subclass + three deployment use cases

**Security requirements** Design goals linked to the attacks they mitigate and the defenses that fulfil them

*Methodology: STORE (Security Threat-Oriented Requirements Engineering) · Taxonomy: NIST AI 100-2e 2025*

SUMMARY

# Progress Against the Expected Results

| | | |
|---|---|---|
| i | **Comprehensive analysis of the state of the art** | Delivered — Paper 1 (under major revision, IJIS); extended by Paper 2's threat & defense analysis |
| ii | **Development of a sustainable real-time AD solution** | Starting — design phase guided by the nine criteria and security requirements |
| iii | **Demonstration on safety-critical CPS** | Future work |
| iv | **Evaluation with sustainability metrics** | Future work |
| v | **Dissemination** | Ongoing — 2 core papers, 3 collaborative papers, summer-school presentation |

# Collaborations and Additional Outputs

## TUAI network collaboration

**Edge-based Safety Optimization: Lightweight Fuzzy Integrated Model for Trustworthy AGV Braking**

Joint work by TUAI doctoral candidates and supervisors across SUT, NTNU and UNINA — grown out of TUAI group activities

## Collaborations within NTNU's Groups

- ICSSPulse: A Modular LLM-Assisted Platform for ICS Penetration Testing — *IFIP Sec, Jun 2026*
- The Emerging Role of Large Language Models in Threat Modelling: A Survey — *IEEE ICMI, Apr 2026*
- Using LLMs to Elicit Security Requirements for Service-Oriented Cyber Ranges — *IEEE CSR CRIRM, 2026*

*These outputs complement DC12 rather than compete with it: they broaden ICS-security and LLM expertise, build co-authoring networks inside TUAI and NTNU.*

# Training and Network Activities

## Scientific training

- Intrusion Detection and Prevention
- Cyber Physical Systems Security
- Research Workshop on Lightweight Models for Real-time Applications

## Transferable skills

- Research Ethics
- Basic University Didactics

## Network & events

- IFIP Sec 2026 – Jun 2026
- TUAI Summer School, Gliwice — Jul 2026

MOBILITY PLAN

# Secondments and Applied Collaboration

### SUT · Gliwice
**4 months**
*[Planned / September-December 2026]*

### UPM · Madrid
**4 months**
*[Unconfirmed / February-May 2027]*

### HVL · Bergen
**4 months**
*[Planned / February–May 2028]*

PLAN TO COMPLETION

# Next Steps

1. **Finalize Paper 1 revision** Submit the revised manuscript to the International Journal of Information Security.
2. **Design the sustainable real-time AD solution** Architecture driven by the nine deployability criteria and the adversarial-robustness security requirements (expected result ii)
3. **Implement & demonstrate** Safety-critical CPS demonstration (expected result iii)
4. **Evaluate with sustainability metrics** Criteria-based evaluation protocol (expected results iv & v)
5. **Thesis integration**

# Thank you — Questions?

Athanasia Kollarou · athanasia.kollarou@ntnu.no

Funded by the European Union's HORIZON TMA MSCA Doctoral Networks · GA No 101168344 (TUAI)







Photo – first slide and this slide: Gløshaugen – Photo: Mentz Indergaard/NTNU
Gjøvik – Photo: Espen Taftø Vestad/NTNU | Ålesund – Photo: Sissel Basso/NTNU