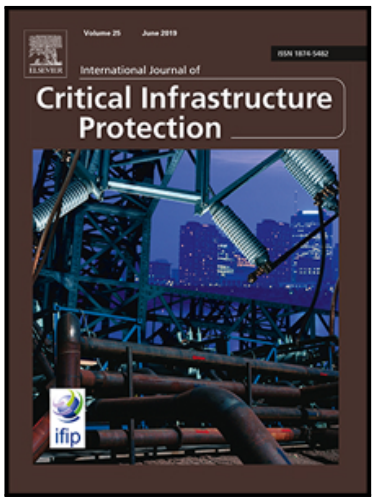


Adaptive Machine Learning based Cyber Threat Intelligence in Industrial Control Systems: A Systematic Literature Review

Nabeel Nasir , Raymond Lutui , Nurul I. Sarkar , Taniela Vaipulu

PII: S1874-5482(26)00046-6
DOI: https://doi.org/10.1016/j.ijcip.2026.100874
Reference: IJCIP 100874

To appear in: *International Journal of Critical Infrastructure Protection*

Received date: 21 November 2025
Revised date: 21 April 2026
Accepted date: 1 July 2026

Please cite this article as: Nabeel Nasir , Raymond Lutui , Nurul I. Sarkar , Taniela Vaipulu , Adaptive Machine Learning based Cyber Threat Intelligence in Industrial Control Systems: A Systematic Literature Review, *International Journal of Critical Infrastructure Protection* (2026), doi: https://doi.org/10.1016/j.ijcip.2026.100874

Adaptive Machine Learning based Cyber Threat Intelligence in Industrial Control Systems: A Systematic Literature Review

Order of Authors: Nabeel Nasir

Raymond Lutui

Nurul I.Sarkar

Taniela Vaipulu

Corresponding Author: Nabeel XXXXX17592 Nasir Auckland University of Technology NEW ZEALAND

Keywords: Cyber Threat Intelligence (CTI), Industrial Control Systems (ICS), Machine Learning, Cross-Layer Defense, Purdue Model, Zero Trust, Adaptive Security, Threat Correlation.

Abstract*—The convergence of information and operational technology has rendered Industrial Control Systems (ICS) prime targets for sophisticated, multi-stage cyber attacks that exploit vulnerabilities across different layers of the industrial enterprise. Defending these critical systems necessitates a paradigm shift from static, isolated security controls to intelligent, adaptive, and holistic approaches. This Systematic Literature Review (SLR) synthesizes and critically analyzes 34 recent academic papers published between 2020-2025 to establish the foundational knowledge required to develop novel, adaptive Cyber Threat Intelligence (CTI) mechanisms that leverage Machine Learning (ML) for cross-layer defense. The review is guided by a central research goal: to understand how to build a unified, cross-layer CTI system capable of correlating threat data, managing performance impact, automating data translation, and overcoming significant adoption barriers. Our extensive analysis reveals that while numerous studies have proposed ML-based CTI and cross-layer defense mechanisms, no single existing approach fully integrates these concepts into a cohesive, adaptive system. Key findings highlight the promise of Deep Reinforcement Learning for adaptive defense orchestration, the critical performance efficiencies of lightweight models for edge deployment, the recognized importance of standardization (e.g., STIX/TAXII) for interoperability, and the persistent barriers of data scarcity, trust deficits, and the simulation-to-reality gap. Based on this comprehensive synthesis, we propose a detailed conceptual multi-tier architectural approach that addresses identified gaps and provides a blueprint for future research. This review concludes by outlining an extensive and actionable research agenda for realizing a truly adaptive, resilient, and intelligent cross-layer CTI system for next-generation ICS security. The findings contribute to both academic research and industrial practice by providing a roadmap for developing next-generation security solutions that can adapt to the evolving threat landscape while maintaining operational integrity in critical infrastructure systems.

***Index Terms*—Cyber Threat Intelligence (CTI), Industrial Control Systems (ICS), Machine Learning, Cross-Layer Defense, Interoperability, Purdue Model, Zero Trust, Adaptive Security, Threat Correlation.**

I. Introduction

A. The Evolving Threat Landscape for Critical Infrastructure

The convergence of Operational Technology (OT) with Information Technology (IT) has transitioned Industrial Con-trol Systems (ICS) from air-gapped, isolated networks to interconnected cyber-physical systems [1]. While this shift towards Industry 4.0 and IIoT improves efficiency, it introduces unprecedented vulnerabilities. High-profile attacks—ranging from Stuxnet's physical destruction to the 2015 Ukraine power grid outage and the Colonial Pipeline ransomware incident—demonstrate that modern adversaries are multi-stage and persistent, capable of causing cascading societal impacts [13]. These incidents, summarized in Table I, underscore the

transition of ICS security from a technical niche to a critical matter of national security.

B. ML-based CTI Landscape and Cross-Layer Defense Foun-dations

To synthesize the current state of adaptive security, it is essential to first introduce background context on (i) how Machine Learning (ML) is currently used across the Cyber Threat Intelligence (CTI) lifecycle, and (ii) what the core operational steps are in cross-layer defense for ICS. This preliminary framing helps explain why our systematic review is structured around four research questions that follow an end-to-end pipeline.

What are current categories for ML-based CTI? Based on our analysis and established CTI standards (e.g., STIX/TAXII) [8], [9], we categorize ML-based CTI into three complementary categories:

1) **Indicator-Centric CTI (IoC Processing and Triage):** ML models classify and prioritize known indicators (e.g., hashes, domains, IPs) and automate ingestion and de-duplication of static intelligence feeds.
2) **Behavior-Centric CTI (TTP and Anomaly In-ference):** ML is used to infer behaviors and tac-tics/techniques/procedures (TTPs) from telemetry (e.g., protocol misuse, sequence anomalies, and process devi-ations), aligning with structured behavioral knowledge bases such as ATT&CK for ICS [38].
3) **Anticipatory CTI (Prediction and Proactive Harden-ing):** ML supports forecasting and prioritization (e.g., likely attack progression, likely next-step TTPs), enabling proactive policy hardening rather than purely reactive response.

What are the core steps for cross-layer defense? Across the Purdue layers, effective cross-layer defense requires a repeatable operational workflow that connects enterprise CTI to OT enforcement under safety and real-time constraints [36], [37]. We summarize the core steps as:

1) **Multi-layer sensing and telemetry:** Collect events across Levels 0–5 (process signals, PLC/HMI logs, network flows, and enterprise alerts).
2) **Normalization and enrichment:** Convert heterogeneous telemetry and CTI artifacts into structured representations (e.g., STIX objects, asset context, protocol semantics).
3) **Cross-layer correlation:** Link indicators and behaviors across layers and time (e.g., IT compromise to OT manipulation), producing a coherent hypothesis of attacker intent.

4) **Decision and orchestration:** Select safe response policies (e.g., monitor, restrict, isolate, verify) consistent with operational constraints and safety envelopes.
5) **Enforcement and feedback:** Deploy device- and protocol-specific actions and continuously learn from outcomes to improve future detection and response.

This framing directly motivates our research questions: SLR-RQ1 focuses on *integration and correlation* strategies (Steps 1–3), SLR-RQ2 evaluates whether the pipeline can run within *latency/computation/network* constraints (Steps 1–5), SLR-RQ3 targets the *semantic translation* required to turn CTI into enforceable OT actions (Steps 2–5), and SLR-RQ4 addresses *adoption* by ensuring explainability, governance, and trust for operators and organizations.

C. The Critical Role of Intelligence-Driven Defense

Traditional cybersecurity approaches, developed primarily for IT environments, prove inadequate when applied to ICS environments. Signature-based detection systems struggle with the deterministic, protocol-specific communications in OT networks. Perimeter-based defenses become ineffective when the perimeter itself becomes blurred through IT/OT convergence. Most critically, reactive security approaches that focus on incident response after a breach has occurred are fundamentally incompatible with environments where a successful attack can cause physical damage, environmental harm, or loss of life.

This reality necessitates a fundamental shift toward intelligence-driven, proactive defense mechanisms. Cyber Threat Intelligence (CTI) represents a mature approach to understanding adversary behaviors, tactics, techniques, and procedures (TTPs). In its most advanced form, CTI transcends simple indicators of compromise (IoCs) such as malicious IP addresses or file hashes. Modern CTI frameworks, exemplified by the MITRE ATT&CK framework and standardized through formats like STIX (Structured Threat Information eXpression), focus on capturing and sharing knowledge about adversary behaviors and methods in a structured, machine-readable format [3].

D. Research Motivation and Problem Statement

How can machine learning-based cyber threat intelligence be integrated with cross-layer defense mechanisms to create an adaptive system that improves threat correlation and automated response across different ICS architectural layers while managing performance impact and overcoming adoption barriers?

This central goal is decomposed into four research ques-tions, each addressing a critical pillar of the cross-layer CTI foundations:

- **SLR-RQ1 (Integration):** Focuses on the unsolved sys-tems problem of bridging intelligence and correlation across fragmented layers.
- **SLR-RQ2 (Performance):** Investigates whether cross-layer ML can operate within the strict real-time timing and resource limits of ICS hardware.

- **SLR-RQ3 (Interoperability):** Addresses the automated semantic translation from high-level CTI standards (e.g., STIX) to low-level OT controls.
- **SLR-RQ4 (Adoption):** Examines the technical and organizational barriers, emphasizing the role of trust, explainability, and policy fit.

These questions are summarized in Table II.

These four research questions form a progressive pipeline: SLR-RQ1 addresses architectural integration, SLR-RQ2 eval-uates deployment feasibility, SLR-RQ3 enables actionable intelligence translation, and SLR-RQ4 ensures real-world adoption through trust and governance. Together, they define the necessary conditions for a fully operational cross-layer CTI system.

E. Research Contributions and Significance

Unlike existing surveys, this work bridges the gap between intelligence-driven CTI and operational cross-layer ICS defense by introducing a unified analytical framework that integrates semantic translation, performance constraints, and trust-aware deployment. This systematic literature review makes several significant contributions to the field of ICS cybersecurity and adaptive threat intelligence:

Comprehensive Knowledge Synthesis: We provide the first systematic analysis specifically focused on the intersection of ML-based CTI and cross-layer ICS defense, synthesizing findings from 34 carefully selected state-of-the-art papers pub-lished between 2020-2025. This synthesis goes beyond simple categorization to identify patterns, gaps, and opportunities for integration.

Gap Analysis and Research Roadmap: Through systematic analysis of the literature against our research questions, we identify critical gaps in current research and practice. Most significantly, we document the absence of integrated approaches that combine automated CTI processing with adaptive cross-layer defense mechanisms.

Performance and Practical Considerations: Our review provides the first systematic analysis of performance consider-ations for ML-based security in ICS, including detailed exami-nation of computational requirements, latency constraints, and deployment strategies for resource-constrained environments.

Adoption Barrier Assessment: Our review includes sys-tematic analysis of both technical and non-technical barriers to adoption, providing insights for researchers, practitioners, and policymakers working to advance ICS security.

Future Research Agenda: We conclude with a detailed, actionable research agenda that prioritizes the most critical areas for future investigation based on our systematic analysis of current capabilities and limitations.

II. Background and Related Work

A. Industrial Control Systems Architecture and the Purdue Model

Understanding cross-layer threats in industrial environments requires an appreciation of the hierarchical structure of modern industrial control systems. The Purdue Enterprise Reference Architecture [1], [36] is the de facto standard for ICS network

TABLE I: Analysis of Seminal Cyber Attacks on ICS: Evolving Vectors and Architectural Impacts

Attack	Year	Sector	Primary Target	Significance and Impact
Stuxnet	2010	Nuclear	Siemens S7-300 PLCs	First known cyber weapon causing physical damage; covertly altered control logic while masking operator visibility.
Ukraine Power Grid	2015	Energy	SCADA / Circuit Breakers	First confirmed cyber-induced power outage; multi-stage attack using phishing, malware, and DDoS.
TRITON (TRISIS)	2017	Petrochemical	Safety Instrumented Systems	Targeted safety controllers directly, posing risks of catastrophic physical damage and loss of life.
Colonial Pipeline	2021	Oil & Gas	Enterprise IT Systems	Ransomware attack on IT systems led to OT shutdown, exposing critical IT/OT interdependencies and economic impact.

TABLE II: Research Questions and Objectives: Mapping the Review to the Cross-Layer CTI Pipeline

RQ	Objective (What We Study)	Why It Matters for ICS (Operational Lens)	Pipeline Steps
SLR-RQ1 (Integra-tion)	Integration strategies for cross-layer sensing and correlation across Purdue levels; architectural patterns and data fusion mechanisms.	ICS attacks are multi-stage and cross boundaries; without cross-layer corre-lation, isolated alerts do not explain attacker intent or process risk.	Sensing, Enrichment, Correlation
SLR-RQ2 (Performance)	Performance overhead and feasibility of ML-driven defense under real-time and resource constraints; emphasis on latency, computa-tion, and network overhead.	Control loops and fieldbuses are timing-sensitive; excessive overhead can destabi-lize processes or prevent timely response.	All steps, especially Orchestration and Enforcement
SLR-RQ3 (Interoperability)	Automated semantic translation from CTI standards (e.g., STIX/TAXII) into OT-relevant, device- and protocol-specific ac-tions.	CTI is often IT-centric; operational value requires deterministic mapping into enforcement mechanisms (policies, rules, PLC safeguards).	Enrichment, Orchestration, Enforcement
SLR-RQ4 (Adoption)	Adoption barriers (technical, organizational, trust); role of explainability (XAI), human-in-the-loop, and governance.	Operators must trust and verify responses in safety-critical environments; gover-nance and incentives shape CTI sharing and deployment.	Feedback, Governance

segmentation, organizing networks into distinct levels (0–5). A detailed security analysis of these layers is provided in Table III.

Unlike traditional IT environments, ICS systems operate under strict real-time, safety-critical constraints, deterministic communication protocols, and limited computational resources. These characteristics fundamentally alter how ML-based CTI systems must be designed, evaluated, and deployed. Rather than isolating layers, modern CTI defense requires holistic, cross-layer coordination across these boundaries to detect and mitigate multi-stage attacks effectively, particularly where IT-level compromises pivot into OT-level process disruption.

B. Evolution of Cyber Threat Intelligence Platforms

Modern CTI frameworks capture knowledge about adversary tactics, techniques, and procedures (TTPs) in structured formats (e.g., STIX/TAXII and MITRE ATT&CK). As illustrated by the Pyramid of Pain (Table IV), which is adapted from the foundational concept proposed by David Bianco [6], effectively disrupting adversaries requires analyzing complex behaviors (TTPs) rather than relying on trivial indicators. Importantly, this pyramid is *not* proposed by the authors of this paper; we include the adapted model

as an analytical lens to structure discussion of CTI “actionability” in ICS. In the context of ICS, this shift is critical: while blocking an IP address is trivial for an attacker to bypass, disrupting a TTP that targets a specific PLC

control logic requires the adversary to fundamentally reinvent their attack chain. The inclusion of this model in our review serves to highlight that most existing ICS security solutions remain at the 'Trivial' to 'Simple' levels, failing to leverage behavioral intelligence for cross-layer defense.

1) Applicability of ML Paradigms to ICS: As summa-rized in Table V, which provides a comparative synthesis of standard ML approaches, supervised techniques are well-suited for classifying known threats, while unsupervised and reinforcement learning play critical roles in anomaly detection and adaptive defense, respectively. Notably, semi-supervised learning emerges as a vital paradigm for ICS [16] due to the extreme scarcity of labeled attack data in critical infrastructure. Selecting the appropriate paradigm requires balancing detection capability against operational overhead and sample efficiency; for instance, while DRL offers maximum adaptability, its "black-box" nature often conflicts with OT safety requirements unless constrained by formal safety layers.

C. Cross-Layer Defense Mechanisms

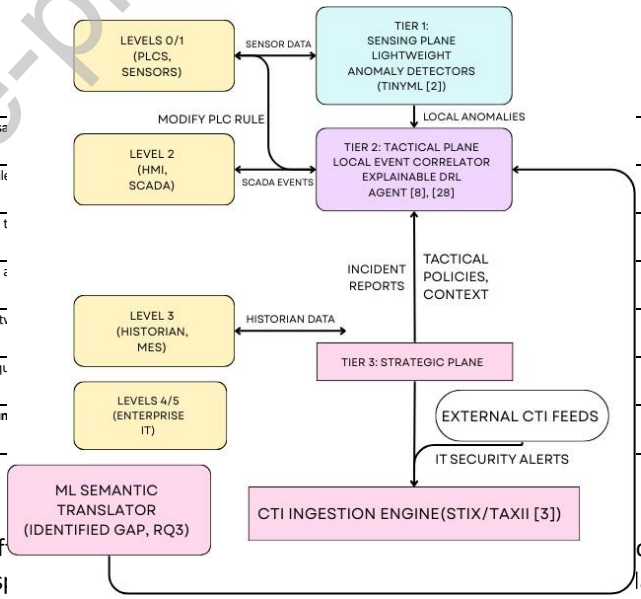
Traditional cybersecurity approaches often focus on indi-vidual system layers or specific types of attacks. However, the sophisticated, multi-stage nature of modern ICS attacks necessitates defense mechanisms that can coordinate across multiple architectural layers and correlate events that may be separated in time and space.

TABLE III: Security Vulnerability Analysis across Purdue Enterprise Reference Architecture Levels

Level	Layer Name	Key Components and Systems	Primary Security Concern
Level 5	Enterprise Network	Corporate IT infrastructure, Internet access, cloud services	Initial point of compromise, external attacker entry point.
Level 4	Enterprise Integration	Business planning, logistics, ERP systems, mail servers	Pivot point for attackers moving from IT to OT environments.
Level 3	Operations Management	Historians, Manufacturing Execution Systems (MES), SCADA servers	Manipulation of operational data, staging point for lower-level attacks.
Level 2	Supervisory Control	Human-Machine Interfaces (HMIs), engineering workstations	Unauthorized control actions, manipulation of operator views.
Level 1	Basic Control	Programmable Logic Controllers (PLCs), Remote Terminal Units (RTUs)	Manipulation of control logic, sending mali-cious commands.
Level 0	Physical Process	Sensors, actuators, valves, pumps, motors	Physical damage, process disruption, and safety incidents.

TABLE IV: The Pyramid of Pain: Analyzing CTI Indicators by Adversary Evasion Cost (Adapted from [6])

Level (from top)	Indicator Type	Impact on Adversa
Trivial	Hash Values	Attacker re-compile
Easy	IP Addresses	Attacker switches t
Simple	Domain Names	Attacker registers a
Moderate	Network/Host Artifacts	Attacker alters net
Challenging	Tools	Attacker must acq
Tough	**TTPs**	**Adversary must fun**



Traditional defense-in-depth strategies rely on multiple, of dern sophisticated attacks necessitate *coordinated and adaptive* res layer defense" to mean a closed-loop workflow that: (i) senses across Purdue levels; (ii) correlates CTI and telemetry into a unified hypothesis; (iii) selects a response policy consistent with operational constraints; and (iv) enforces actions at the device/protocol level with feedback for continuous improve-ment [36], [37]. Fig. 1 visualizes this end-to-end flow and highlights the role of semantic translation as the key bridge between CTI and OT enforcement. This foundation explains why integration (SLR-RQ1) is a first-order problem: without cross-layer correlation and orchestration, high-level CTI cannot reliably prevent lateral movement or process manipulation.

D. Comparison with Existing Surveys and Claimed Novelty

The field of ICS cybersecurity has been the subject of several comprehensive surveys, each focusing on different aspects of the domain. However, a significant gap remains at the intersection of intelligence-driven defense and cross-layer operationalization. Table VI provides a structured comparison of this review against the most relevant existing literature, highlighting our unique focus on semantic translation and operator trust in the context of cross-layer CTI.

What is new beyond existing surveys? Unlike literature that treats CTI and ICS machine learning as isolated domains, this review rigorously maps their intersection. Specifically, our contributions include:

- **Pipeline-Driven Problem Decomposition:** We frame the domain as an end-to-end cross-layer CTI pipeline and map evidence to four research questions that correspond to operational steps rather than isolated model categories.

Fig. 1: Synthesized Architectural Flow (Cross-Layer CTI Pipeline): The figure summarizes the end-to-end workflow that this review argues is missing in prior art, connecting CTI ingestion (STIX/TAXII) to cross-layer correlation, semantic translation, and safe OT enforcement. It is intended as an analytical bridge between SLR-RQ1 (integration), SLR-RQ2 (feasibility under constraints), SLR-RQ3 (CTI-to-action trans-lation), and SLR-RQ4 (trust and governance).

- **Evidence-Weighted Critique:** Beyond descriptive sum-maries, we consistently extract strengths, weaknesses, and operational failure modes (e.g., semantic translation, deployment feasibility, trust) and relate them to ICS constraints.
- **Operationalization of Semantic Translation:** We syn-thesize an actionable, layered translation pipeline that clarifies required inputs/outputs from CTI standards down to OT enforcement artifacts.
- **Trust and XAI Taxonomy for OT:** We provide an ICS-specific taxonomy for explainability and trust mechanisms and connect it to adoption barriers and human-in-the-loop requirements.

TABLE V: Critical Performance and Implementation Trade-offs of ML Paradigms in ICS Environments (Synthesized from [1], [5], [16], [31])

ML Paradigm	Description	Advantages in ICS Context	Challenges and Limitations in ICS Context
Supervised Learning	Learns from labeled data (e.g., "at-tack" vs. "normal") to make predic-tions [5].	High accuracy when trained on representative data.	Scarcity of labeled attack data in ICS environments.
Unsupervised Learn-ing	Identifies patterns and anomalies in unlabeled data [1].	Ideal for zero-day threat detection; no labeled data needed.	High false positive rates due to opera-tional noise.
Semi-Supervised	Combines small labeled datasets with large unlabeled sets [16].	Addresses the 'Small Data' problem in critical infrastructure.	Requires high-quality initial labels for propagation.
Reinforcement Learning	Agent learns optimal policy through environment interaction [31].	Enables fully adaptive and auto-mated defense orchestration.	Safety concerns during the trial-and-error learning phase.

TABLE VI: Comparative Analysis: Positioning This SLR against Existing Literature

Survey	Scope	ML	CTI	ICS/OT	Cross-Layer	Semantic Trans.	XAI/Trust	Our Novelty
Ko et al. [1]	ML for ICS	✓	×	✓	×	×	Partial	Baseline ML application
Aldhaheri et al. [4]	Deep Learning IoT	✓	×	Partial	×	×	×	IoT-focused Deep Learning
Various CTI Surveys	CTI Frameworks	Partial	✓	×	×	×	×	IT-centric intelligence
This Review	**Adaptive CTI/ICS**	✓	✓	✓	✓	✓	✓	**Integrated Pipeline**

- **Integrated Architectural Synthesis:** We propose a conceptual multi-tier architecture that unifies CTI inges-tion, correlation, translation, and adaptive response under performance constraints.

The Central Gap: The core deficiency identified across the literature is the absence of a closed-loop system that connects CTI ingestion, cross-layer correlation, semantic translation, and adaptive response under real-world ICS operational constraints.

III. Systematic Review Methodology

This systematic literature review follows established best practices for rigorous and reproducible research synthesis. The methodology is guided by the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) framework, adapted to the requirements of computer science and cyberse-curity research.

A. Review Protocol and Research Questions

A detailed review protocol was defined prior to the literature search, specifying research objectives, search strategy, selection criteria, and analytical methods. This ensured systematic coverage of the domain while maintaining alignment with the study objectives.

The overarching goal of this review is to examine how integrated ML-based CTI systems can support cross-layer defense in ICS environments. This goal is addressed through four research questions:

SLR-RQ1 analyzes architectural approaches for cross-layer correlation and integration.

SLR-RQ2 investigates the performance impact of ML-based security mechanisms under ICS operational constraints.

SLR-RQ3 examines automation and interoperability chal-lenges, focusing on translating high-level CTI into device-specific actions.

SLR-RQ4 explores adoption barriers, including organiza-tional, economic, and trust-related constraints.

B. Search Strategy and Information Sources

1) Database Selection: The literature search was conducted across four major academic databases to ensure comprehensive coverage:

IEEE Xplore for ICS, industrial automation, and cyberse-curity research.

ACM Digital Library for computer science, machine learning, and security research.

Scopus for multidisciplinary coverage across engineering and applied sciences.

ScienceDirect for high-impact journals in engineering and computer science.

2) Supplementary Search Strategies: To enhance coverage, additional strategies were employed:

Backward Citation Analysis to identify foundational studies.

Forward Citation Analysis using Google Scholar to capture recent developments.

Gray Literature Search targeting major cybersecurity conferences and technical repositories.

Expert Consultation to ensure inclusion of influential research and key contributors.

C. Study Selection Process

1) Inclusion and Exclusion Criteria: Prior to the literature search, specific criteria were established. To maintain academic rigor, the following inclusion criteria were applied: (i) studies published between 2020–2025; (ii) primary focus on ML-based security in ICS or CTI frameworks; (iii) presence of empirical or technical validation. Crucially, while non-peer-reviewed materials (e.g., preprints, blog posts, white papers) were excluded, high-quality peer-reviewed conference papers were explicitly included to capture state-of-the-art developments that frequently emerge in top-tier security venues (e.g., IEEE S&P, USENIX Security, ACM CCS) before journal publication.

2) Selection Process: The selection process comprised four stages:

Stage 1: Initial Screening — Independent title and abstract review by two researchers, with disagreements resolved through discussion.

Stage 2: Full-Text Review — Detailed evaluation of methodology, technical depth, and relevance.

Stage 3: Quality Assessment — Structured assessment of rigor, validation quality, and contribution significance.

Stage 4: Citation Review — Forward and backward citation analysis to identify additional relevant studies.

D. Data Extraction Framework

1) Structured Data Extraction: A structured data extraction framework was developed to ensure consistent analysis of selected studies. Each paper was examined using standardized forms capturing bibliometric details and technical content aligned with the research questions.

Extracted data included:

Bibliometric Data: Publication details, affiliations, geo-graphical distribution, and funding information.

Technical Content: Problem scope, ML techniques, ICS layers addressed, validation methods, performance metrics, and reported limitations.

RQ-Specific Contributions: Architectural integration (RQ1), performance metrics (RQ2), interoperability mechanisms (RQ3), and adoption barriers (RQ4).

Table VII summarizes the data extraction categories.

TABLE VII: Data Extraction Categories and Operationalization

Category	Extracted Fields (Examples)
Bibliometrics	Year, venue (journal/conference), application domain, geographic context (if stated).
ICS Scope	Purdue levels covered, protocols (e.g., Modbus/DNP3), assets considered (PLC/HMI/historian), threat model assumptions.
ML Technique	Paradigm (supervised/unsupervised/semi-supervised/RL), model family (e.g., GNN, Transformer), training setting (centralized/federated), data requirements.
Validation	Dataset/testbed type, realism (simulated vs. high-fidelity), reproducibility signals (hyper-parameters, code/data availability).
Evaluation Metrics	Performance overhead (latency, computation, network), detection/response metrics (accuracy, false positives), safety/operational considera-tions (if evaluated).
RQ Mapping	Evidence relevant to SLR-RQ1–RQ4 (integra-tion, performance, interoperability, adoption) and stated limitations/gaps.

2) Quality Assessment Framework: To ensure the robustness of our synthesis, we utilized a quantified 10-point Quality Assessment (QA) scoring system. Each study was independently evaluated by two researchers across five dimensions, as detailed in Table VIII. Each dimension was scored from 0 to 2: 0 (not addressed), 1 (partially addressed or limited detail), and 2 (rigorously addressed and validated).

The cumulative QA score was computed as the sum of the five dimensions (maximum 10 points). Only studies achieving a cumulative QA score of ≥ 7 (out of 10) were included for

TABLE VIII: Systematic Literature Review Quality Assessment (QA) Scoring Criteria

QA Dimension	Measurement Criterion
Methodological Rigor	Is the study design clearly defined with appropriate ML metrics?
Technical Novelty	Does the work propose a novel model, architecture, or optimization?
Reproducibility	Are datasets, hyper-parameters, and experimental setups provided?
Industrial Relevance	Does the study address realistic ICS protocols (Modbus, DNP3, etc.)?
Validation Quality	Is the model validated on high-fidelity testbeds or realistic datasets?

full thematic synthesis. Studies meeting the primary inclusion criteria but scoring below 7 were retained for background context only, ensuring that our primary analytical findings are based on the most robust evidence available.

To address evidence weighting explicitly, we use QA as an *evidence-strength tiering* mechanism throughout the synthesis: claims about operational feasibility, cross-layer integration, and automation are treated as strongest when supported by studies with high validation quality and industrial relevance, while lower-scoring studies are used primarily to motivate research gaps and to surface under-explored ideas rather than to substantiate definitive conclusions.

E. Analytical Approach

1) Thematic Analysis: A systematic thematic analysis was applied using the following steps:

Initial Coding aligned with research questions.

Pattern Identification across studies.

Thematic Synthesis into higher-level findings.

Gap Analysis to identify unresolved challenges and research opportunities.

2) Quantitative Analysis: Where applicable, quantitative analysis was conducted, including:

Publication Trends across time and venues. **Methodological Distribution** of ML techniques. **Performance Metric Comparison** across comparable stud-

ies.

Citation Analysis to identify influential research and authors.

F. Ensuring Rigor and Minimizing Bias

1) Multiple Reviewer Approach: All critical review stages involved multiple independent reviewers. Disagreements were resolved through documented discussion to minimize bias.

2) Protocol Pre-registration: The review protocol was finalized prior to the search process, reducing the risk of post-hoc bias.

3) Transparency and Documentation: Detailed documenta-tion—including search logs, selection decisions, and analysis records—was maintained to support reproducibility and inde-pendent verification.

IV. Results: Comprehensive Literature Analysis

Our systematic analysis of 34 carefully selected papers reveals a rich but fragmented research landscape. This section

presents findings organized around our four research questions and emphasizes critical comparison rather than tutorial-style description: for representative studies we extract strengths, weaknesses, and operational failure modes (e.g., interoperability, safety constraints, and deployability), and then synthesize cross-study patterns into evidence-backed gaps and directions.

A. Overview of Selected Literature

1) Publication and Geographic Distribution: The selected papers span a diverse range of publication venues, reflecting the interdisciplinary nature of this research domain. IEEE trans-actions and conferences account for 47% of selected papers, with particularly strong representation in IEEE Transactions on Industrial Informatics, IEEE Internet of Things Journal, and IEEE Transactions on Dependable and Secure Computing. Computer science venues, including ACM conferences and Elsevier journals, account for 35% of publications. The remaining papers come from specialized security conferences and interdisciplinary venues.

2) Temporal Distribution and Research Evolution: Publica-tion patterns show clear evolution in research focus over the review period. Early papers (2020-2021) primarily focused on applying existing ML techniques to ICS security challenges, often treating threat intelligence and defense mechanisms as separate problems. The middle period (2022-2023) saw increasing attention to integration challenges and cross-layer approaches. Recent papers (2024-2025) demonstrate more sophisticated understanding of the unique requirements for adaptive, integrated systems.

3) Methodological Approaches: The methodological ap-proaches employed across the selected papers reveal interesting patterns. Experimental evaluation using simulated data or testbeds accounts for 62% of papers, reflecting the challenges of obtaining real-world ICS attack data. Theoretical analysis and architectural proposals account for 21% of papers, while field studies and case analyses account for the remaining 17%. This distribution highlights both the maturity of simulation-based research approaches and the ongoing challenges in bridging from laboratory research to real-world deployment.

B. SLR-RQ1: Architectural Integration and Cross-Layer Cor-relation

1) Current State of CTI Architecture Research: Existing literature shows substantial interest in cyber threat intelligence (CTI) architectures, though with varying degrees of integration and operational relevance. At a foundational level, several works focus on CTI data collection, processing, and sharing mechanisms.

Haque et al. [3] propose an automated CTI-sharing platform combining NLP and ML for STIX extraction. **Strengths:** Highly effective at automated intelligence ingestion and nor-malization across varied feeds. **Weaknesses:** It critically fails to bridge the IT-OT gap, offering zero operational mechanisms for how these STIX rules alter ICS PLCs.

Alazab et al. [23] extend this by introducing multi-source intelligence fusion. **Strengths:** Achieves high accuracy in network-level correlation and expands contextual awareness

natively. **Weaknesses:** The framework remains inherently network-centric (Levels 3-4) and fails to bridge the IT-OT gap due to a lack of specialized semantic mapping between CTI formats and OT control protocols. This is critical because network-level alerts without device-level enforcement cannot prevent localized process manipulation.

Alajmi et al. [11] present an automated deep learning-based threat detection approach for cyber-physical environments. **Strengths:** Demonstrates strong detection performance via optimized deep models. **Weaknesses:** Serves largely as an isolated detection engine, lacking integration with a reactive orchestration architecture. This is critical because high accuracy in detection is operationally insufficient if it does not trigger an automated, safe containment action.

2) Critical Synthesis of Architectural Research: Across the reviewed studies, a clear pattern emerges: integration remains largely vertical or siloed. While systems like Haque et al.

[3] and Alazab et al. [23] excel at enterprise-level (L4–L5) intelligence processing or network-level (L3) detection, they fail to achieve true cross-layer coordination with edge level (L0–L1) operations. This failure is often due to the *semantic translation gap*–the inability to map high-level intelligence to deterministic industrial protocols. Furthermore, while DRL approaches (e.g., Wilson et al. [31]) offer a path toward autonomous defense, their lack of explainability can prevent them from being trusted for automated containment in safety-critical loops. Most contemporary literature achieving high detection performance (e.g., Alajmi et al. [11]) operates as an isolated detection engine rather than a closed-loop, CTI-driven orchestration system.

3) Transition to Operational Constraints: While the architec-tural research provides the "what" and "where" of integration, it frequently overlooks the "how fast." The next research question (SLR-RQ2) investigates whether these proposed integrations are actually deployable under the strict performance envelopes of industrial hardware.

4) Integration Gaps and Reasoning: To avoid purely list-style reporting, we identify gaps using an explicit analytical lens grounded in the cross-layer defense workflow (sensing, enrichment, correlation, orchestration, enforcement) and in the interoperability and safety constraints reflected in common ICS guidance and standards (e.g., STIX/TAXII, ATT&CK for ICS, and IEC 62443) [8], [9], [37], [38]. Under this lens, a “gap” is recorded when a study either (i) fails to connect one pipeline stage to the next (e.g., CTI ingestion without OT enforcement), or (ii) cannot justify that the connection is feasible and safe under ICS operational constraints (e.g., timing, vendor heterogeneity, and operator trust).

Systematic analysis of the literature highlights several persistent integration gaps, summarized in Table IX. These gaps are not merely technical but rooted in operational reality:

- **Semantic Translation Gap:** Rooted in the mismatch between IT-centric threat schemas (e.g., STIX) and OT-specific control semantics (e.g., PLC register maps).
- **Temporal Coordination Gap:** Arises because CTI typ-ically arrives at strategic timescales (hours/days), while ICS defense requires deterministic responses (millisec-onds/seconds).

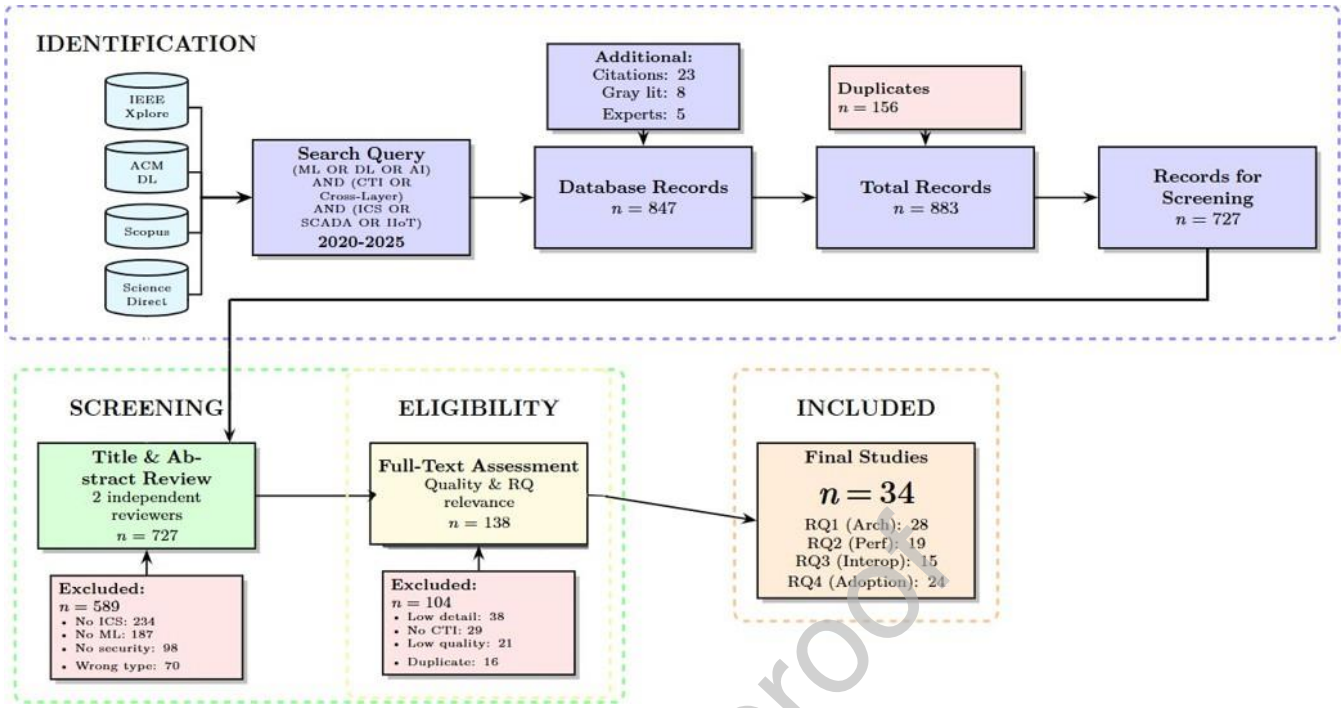


Fig. 2: Synthesis of the 34 Selected Articles: A Critical Mapping of Research Contributions against methodological rigor, operational constraints, and integration gaps. The figure is used as an evidence map (not merely a summary), highlighting where the literature concentrates (e.g., detection accuracy) versus where it remains sparse (e.g., validated end-to-end cross-layer orchestration and CTI-to-action translation).

- **Contextual Integration Gap:** Caused by the historical isolation of security telemetry from process telemetry (His-torians/MES), making multi-layer fusion computationally expensive.

While technical correlation mechanisms are explored, few works address the need for transparent and explainable decision-making required for operator trust in industrial environments.

5) *Emerging Architectural Patterns:* Despite these chal-lenges, several promising architectural patterns are emerging: **Hierarchical Intelligence Processing:** Intelligence is pro-cessed across multiple abstraction levels, aligning with the hierarchical structure of ICS environments.

Federated Defense Coordination: Distributed defense agents operate autonomously at different layers while coordi-nating through structured communication, improving scalability and responsiveness.

Adaptive Policy Orchestration: Security policies are dynamically adjusted based on evolving threat intelligence and operational context, enabling adaptive and context-aware defense.

C. SLR-RQ2: Performance Impact and System Efficiency

The literature demonstrates growing attention to performance considerations for ML-based security in ICS. We prioritized three primary metrics for evaluation—**Latency**, **Computation**, and **Network Overhead**—because they represent the funda-mental operational bottlenecks in deployed ICS:

- **Latency:** Selected because it directly impacts control-loop stability and the ability to intervene before safety limits are breached.
- **Computation:** Selected due to the prevalence of low-power PLC, RTU, and gateway hardware with limited RAM/CPU cycles.
- **Network Overhead:** Selected because fieldbus stability (e.g., Modbus over RS-485) is highly sensitive to band-width congestion from security telemetry.

While secondary metrics such as storage, energy usage, and recall are important, these three are the most critical for maintaining operational continuity.

Alwasi et al. [2] present a detailed performance evaluation of TinyML-based anomaly detection tailored for resource-constrained ICS edge devices. **Strengths:** Provides rigorous quantitative evaluation of computational latency and memory usage on actual hardware (Raspberry Pi), demonstrating sub-40ms inference. **Weaknesses:** The evaluation is localized to anomaly detection and fails to explore how these edge alerts fuse with higher-layer CTI due to the lack of a standardized cross-layer messaging protocol. This is critical because isolated edge alerts cannot contribute to a proactive, system-wide defense posture without effective escalation.

Zhang et al. [27] extend this work by analyzing compressed ML models. **Strengths:** Achieves a significant 92% reduction in model size while maintaining high accuracy (> 97%), proving the viability of model pruning for ICS. **Weaknesses:** Like Alwasi et al., it focuses on isolated performance rather than the system-wide impact of CTI integration due to a lack of end-to-

TABLE IX: Identified Gaps in Existing ICS CTI Architectures (SLR-RQ1) (Synthesized from [3], [23], [26], [32])

Identified Gap	Challenge Description	Implication for Future Systems
Semantic Translation	Difficulty in mapping abstract CTI (e.g., STIX TTPs) to concrete, device-level defensive actions.	Necessitates ML-based translation engines, ontolo-gies, or graph models to bridge IT–OT semantics.
Temporal Coordination	Disconnect between long-term intelligence analysis and real-time defense orchestration.	Requires correlation of slow-moving indicators with rapid, high-impact events.
Contextual Integration	Limited fusion of security data with operational context, increasing false positives.	Future systems must integrate with MES, historians, and operational workflows.
Explainability and Trust	Black-box ML/DRL decisions lack transparency for operators.	Architectures must embed XAI techniques to justify automated responses.

end architectural testing. This is critical because model pruning may inadvertently remove features required for correlating low-level process anomalies with high-level threat indicators.

1) System-Level Performance Analysis: Several studies move beyond individual model efficiency to evaluate system-wide performance impacts.

Hao et al. [19] propose a hybrid statistical–ML framework. **Strengths:** Successfully combines statistical process control with ML, achieving rapid detection (sub-12s) for multi-stage attacks. **Weaknesses:** The end-to-end latency analysis lacks a detailed breakdown of communication overhead between Levels 1 and 3 due to the use of idealized network models. This is critical because network jitter and congestion on industrial fieldbuses can significantly delay the dissemination of time-sensitive threat intelligence.

In comparison to Hao et al., Chen et al. [25] focus on deep correlation using graph neural networks (GNNs). **Strengths:** Superior cross-layer correlation capability compared to statistical methods. **Weaknesses:** The computational cost is significantly higher ($O(n \log n)$ post-optimization), making it less suitable for Level 1 enforcement compared to the lightweight approaches of Alwasi et al. [2] due to the high memory footprint of graph embeddings. This is critical because high-complexity models can cause processing bottlenecks that lead to buffer overflows and missed control signals in real-time PLC environments.

2) Performance–Security Trade-off Analysis: Managing trade-offs remains a central challenge. Martinez et al. [30] analyze performance-aware ML deployment. **Strengths:** Ex-plicitly models the trade-off between detection accuracy and strict real-time constraints. **Weaknesses:** The analysis is purely theoretical and lacks a proposed automated mechanism to dynamically adjust these trade-offs in real-time due to the absence of a feedback control loop between the ML agent and the process historian. This is critical because static trade-offs cannot adapt to the dynamic threat levels and varying network conditions common in operational industrial environments.

3) Benchmarking and Standardization Gaps: Despite in-creased focus on performance, the literature exposes notable evaluation gaps:

Lack of Standardized Metrics: Studies employ heteroge-neous performance metrics, limiting direct comparison across approaches and hindering reproducibility.

Inconsistent Evaluation Environments: Performance eval-uations span diverse hardware platforms and operational scenarios, complicating generalization of results.

Limited Real-World Validation: Most evaluations rely on simulations, with few studies validating performance in live industrial environments subject to real operational interference and constraints.

Table X summarizes the dominant performance optimization strategies reported across the reviewed studies, highlighting re-curring mechanisms that aim to preserve OT timing guarantees while enabling deeper analytics when required.

4) Synthesis: Bridging Performance and Intelligence: The findings from SLR-RQ2 highlight a fundamental tension: while edge-level optimizations like TinyML [2] and model pruning

[27] make ML-based security feasible on constrained hardware, the system-wide performance impact of deep cross-layer correlation remains an open challenge. This fails primarily due to the lack of standardized benchmarks for ICS security latency, which prevents a unified comparison of these approaches. This identifies a critical need for transition—from isolated model optimization to automated, interoperable translation mechanisms that can scale intelligence without overwhelming fieldbus capacity. Achieving this balance is essential because even perfect detection is useless if the system cannot respond within the physical process's safety-critical time constant.

D. SLR-RQ3: Automated Translation and Interoperability

1) The Standardization Landscape: The literature consis-tently highlights the importance of standardization for inter-operability in ICS environments, while also emphasizing the practical challenges posed by system heterogeneity.

Among existing standards, the STIX/TAXII framework is the most widely adopted for threat intelligence representation and sharing. Haque et al. [3] demonstrate an advanced STIX-based CTI processing system capable of automatically extracting and structuring intelligence from more than 20 heterogeneous sources. Their work illustrates how IoCs, TTPs, and contextual metadata can be normalized into standardized threat represen-tations, improving interoperability at the intelligence-sharing level.

Rodriguez et al. [26] address interoperability from a policy perspective by proposing an adaptive Zero Trust architecture with standardized security policy templates. These templates can be automatically translated into vendor-specific config-urations for different ICS devices. Although their approach does not fully resolve semantic translation, it demonstrates the feasibility of standardized policy abstraction for certain control mechanisms.

TABLE X: Performance Optimization Strategies for ML in ICS (SLR-RQ2) (Synthesized from [2], [27], [30], [33])

Strategy	Mechanism and Rationale	Primary Benefit
Hierarchical Processing	Lightweight, low-latency analysis (e.g., TinyML) at edge levels (L1–L2), with escalation to higher levels (L3+) for complex correlation.	Rapid detection, reduced network load, and minimal edge overhead.
Adaptive Complexity	Dynamic scaling of computational intensity based on detected threat levels.	Preserves OT performance while enabling deep analysis when required.
Model Compression	Quantization, pruning, and distillation reduce infer-ence cost on constrained devices.	Enables real-time ML inference on low-power ICS hardware.
Distributed Processing	Partitioned analytics and federated learning avoid centralized raw data collection.	Enhances scalability, resilience, and data privacy.

2) The Semantic Translation Challenge: The most critical gap identified in the literature concerns automated semantic translation between high-level CTI and device-level operational controls. This challenge arises across multiple dimensions:

Conceptual Translation: Abstract threat concepts (e.g., lateral movement or privilege escalation) must be mapped to concrete indicators relevant to specific industrial processes. **Protocol Translation:** Network-level threat indicators must be contextualized within industrial protocols such as Modbus, DNP3, or EtherNet/IP, requiring detailed knowledge of protocol semantics and misuse patterns.

Device-Level Translation: Security policies must be con-verted into vendor-specific configurations while accounting for device capabilities and operational constraints.

Taylor et al. [32] provide one of the most comprehensive analyses of cross-vendor interoperability challenges, examining 12 ICS platforms and identifying over 200 incompatibilities in threat representation and response. They propose a multi-layer translation architecture comprising a semantic mapping layer, a protocol translation layer, and a configuration generation layer. However, their work remains largely conceptual, with limited empirical validation.

3) Operationalizing Semantic Translation in Cross-Layer CTI: To move beyond conceptual discussions, we propose an operationalized four-layer pipeline for semantic translation in ICS environments. This architecture addresses the critical "semantic gap" between abstract IT-centric threat intelligence and the deterministic reality of OT control logic.

1) **Threat Representation Layer:** Ingests standardized CTI formats such as STIX 2.1 and TAXII. This layer identifies the core TTPs and IoCs (e.g., "Lateral Movement via SMB" or "Modbus Register Manipulation").
2) **Semantic Enrichment Layer:** Maps abstract TTPs to specific industrial assets, process functions, and protocol roles. For instance, a "Credential Access" TTP is enriched with context regarding the specific HMI or engineering workstation credentials at Level 2.
3) **Control Interpretation Layer:** Converts the enriched intelligence into localized security policy classes, such as *Monitor, Restrict, Isolate, Verify,* or *Rate-limit*. This layer utilizes domain-specific ontologies [24] to ensure policies do not violate process safety constraints.
4) **Device-Level Action Layer:** The final deterministic conversion into protocol-specific actions, such as Python-based automated Modbus register restrictions or PLC block updates.

Despite the promise of this pipeline, our review identifies a persistent research gap: existing literature rarely provides validated, closed-loop mapping from CTI semantics to protocol-level enforcement. Most studies (e.g., Taylor et al. [32]) remain at the conceptual layer, lacking the rule-generation engines required for multi-vendor ICS interoperability.

4) Synthesis: From Standards to Execution: The analysis of SLR-RQ3 reveals that while STIX/TAXII provides a robust foundation for intelligence sharing [3], the automated translation of these representations into operational control actions remains the primary bottleneck. This fails primarily due to the heterogeneous nature of ICS vendor protocols. Solving this matters in real ICS because manual translation by operators introduces unacceptable delays during rapid-onset cyber-physical attacks.

5) Industry and Standards Body Initiatives: Beyond aca-demic research, several works highlight ongoing efforts by industry consortia and standards bodies to improve CTI interoperability. These initiatives aim to align threat intelligence formats, security policies, and operational practices, providing essential foundations for scalable and interoperable ICS security architectures.

E. SLR-RQ4: Adoption Barriers and Trust Mechanisms

The literature extensively documents various technical bar-riers that hinder the adoption of advanced ML-based threat intelligence systems in ICS environments. These barriers span multiple dimensions and require multifaceted solutions.

Table XI provides a synthesized summary of the most fre-quently reported adoption barriers and the mitigation strategies proposed across the reviewed literature.

1) Technical Barriers to Adoption: **Data Scarcity and Quality Issues**: This emerges as perhaps the most frequently cited barrier across the literature. Ferrag et al. [18] address this challenge through the creation of comprehensive, realistic datasets for IIoT security research. Their Edge-IIoTset dataset includes diverse attack scenarios across different industrial protocols and device types, providing a valuable resource for training and validating ML models.

Real-Time Performance Requirements: ICS environments impose strict timing constraints that can conflict with the computational requirements of sophisticated ML algorithms.

2) Explainable AI for ICS Security: Taxonomy and Role in Operator Trust: Trust is frequently cited in the literature as a primary adoption barrier (SLR-RQ4). To transition this from a conceptual concern to an operationalizable framework, we

TABLE XI: Identified Adoption Barriers and Mitigation Strategies (Synthesized from [18], [28], [33], [41], [46])

Category	Barrier	Description and Solutions
Technical	Data Scarcity & Quality	Limited availability of realistic, labeled ICS attack data restricts model training. *Solutions: Realistic datasets [18], federated learning [33], transfer learning.*
	Performance Overhead	Real-time OT constraints conflict with compute-intensive ML models. *Solutions: TinyML, hardware acceleration, adaptive models [2].*
Trust	Lack of Explainability	Black-box decisions reduce operator trust and adoption. *Solutions: XAI frameworks [29], model-agnostic expla-nations, causal analysis.*
	Automated Error Risk	Incorrect automated actions may disrupt operations or safety. *Solutions: Human-in-the-loop control, digital twins [17], confidence scoring.*
Organizational	CTI Sharing Reluctance	Fear of legal, IP, or exposure risks limits information sharing. *Solutions: Anonymized platforms, blockchain-based CTI [28], policy incentives.*
	Cost & ROI	High initial investment with unclear return on security spending. *Solutions: Phased deployment, TCO analysis, cloud/SaaS adoption.*

propose an XAI taxonomy tailored specifically for ICS security environments.

A. Taxonomy by Explanation Timing and Scope:

- **Ante-hoc (Intrinsic):** Utilizing inherently interpretable models (e.g., decision trees or rule-based logic) for Level 1 operations where transparency is a safety requirement.
- **Post-hoc Explanations:** Applying model-agnostic tech-niques like SHAP or LIME [29], [42] to complex deep learning models used for cross-layer correlation.
- **Global vs. Local Scope:** Global explanations provide an audit trail of the entire model's behavior, while local explanations justify specific, high-priority alerts (e.g., "Why was this Modbus command blocked?").

B. Taxonomy by Operational Purpose: The effectiveness of XAI in ICS is measured by its ability to answer four critical operator questions:

- **Alarm Justification:** "Which specific sensor variables or network features triggered this alert?"
- **Cross-Layer Traceability:** "How were enterprise-level indicators correlated with this low-level process anomaly?"
- **Safety Assurance:** "Is the recommended response action (e.g., isolating a PLC) safe for the current physical process state?"
- **Actionability:** "What specific evidence supports the proposed intervention?"

Beyond algorithmic transparency, trust requires integration with HMI interfaces and operator workflows [49]. **Confidence Indicators** must quantify certainty, while **Digital Twin Sim-ulation** [17] allows operators to non-destructively preview the impact of automated responses before confirming them within a **Human-in-the-Loop** framework. Unlike IT systems, incorrect automated decisions in ICS environments can result in physical damage or safety hazards, making explainability and verification essential rather than optional.

3) Organizational and Policy Barriers: Beyond technical challenges, the literature highlights significant organizational and policy barriers that impede adoption.

Reluctance to Share Information: Despite the benefits of collaborative defense, asset owners are often hesitant to

share threat intelligence due to concerns about exposing vulnerabilities, revealing proprietary operational details, or incurring legal liability. Approaches like federated learning [33] and blockchain-based sharing [28] are proposed as technical solutions to mitigate these concerns, but cultural and business barriers remain.

Skills Gap: Implementing and managing sophisticated ML-based security systems requires a unique combination of expertise in OT, IT security, and data science. This skill set is rare, and many organizations struggle to attract and retain the necessary talent.

Cost and ROI: The cost of acquiring, deploying, and maintaining advanced security systems can be substantial. Demonstrating a clear return on investment (ROI) for security expenditures is notoriously difficult, particularly for proactive measures designed to prevent incidents that have not yet occurred.

4) Synthesis: Establishing a Trust-Based Defense Ecosystem: The barriers identified in SLR-RQ4 demonstrate that technical excellence in ML detection is insufficient for ICS adoption. The integration of XAI taxonomies, digital twins [17], and privacy-preserving sharing [28], [33] is essential to reconcile the "black-box" nature of adaptive ML with the risk-averse culture of industrial operations. This matters because without operator trust, even the most accurate CTI system will be bypassed or disabled in high-stakes operational settings.

V. Architectural Synthesis: A Conceptual Approach

Drawing upon the comprehensive analysis of the literature, we propose a conceptual multi-tier architecture designed specifically to address the identified integration gaps and leverage the most promising existing approaches. This architecture, depicted in Fig. 1, provides a blueprint for a unified, adaptive, and intelligent cross-layer CTI system.

A. A Multi-Tier, Multi-Plane Architecture

The proposed architecture is organized into three logical tiers that operate across different timescales and levels of abstraction,

directly addressing the gaps identified in our analysis.

Tier 1: Real-Time Sensing & Enforcement Plane. This tier operates at the edge, closest to the physical process (Levels 0-2). It is composed of lightweight, computationally efficient anomaly detectors deployed on or near ICS devices. The findings of Alwasi et al. [2] and Zhang et al. [27] validate the feasibility of this approach, confirming that TinyML models can achieve high accuracy with minimal performance impact, satisfying the strict requirements identified in SLR-RQ2. This tier's primary role is to provide immediate detection of deviations from normal behavior and feed high-quality, localized alerts to the next tier.

Tier 2: Adaptive Defense & Tactical Plane. This is the intelligent core of the architecture, responsible for correlation and decision-making. A *Local Event Correlator* aggregates alerts from Tier 1 with data from supervisory systems (Levels 2-3) to build a rich, contextual picture of the system's state. This fused information forms the state representation for the *Adaptive DRL Agent*, informed by recent DRL-based defense orchestration approaches [31]. This agent is responsible for learning and executing optimal defense policies that span mul-tiple layers. Crucially, it must be augmented with explainability mechanisms [29] to address the trust deficit identified in SLR-RQ4, providing operators with clear justifications for its actions. This tier directly addresses the need for cross-layer correlation central to our main research goal.

Tier 3: Global Intelligence & Strategic Plane. This tier operates at the enterprise level (Levels 4-5) and serves as the bridge between the local ICS environment and the global threat landscape. It includes a *CTI Ingestion Engine* based on established standards like STIX/TAXII, as demonstrated by Haque et al. [3]. Its most critical and novel component is the *ML-based Semantic Translator*. This component is designed to solve the interoperability and automation gap identified in SLR-RQ3. It takes high-level CTI (e.g., STIX objects describing TTPs) and translates them into actionable, context-aware tactical policies and monitoring rules for the Tier 2 DRL agent and Tier 1 detectors. This translation engine would leverage techniques like graph neural networks [25] or domain-specific ontologies [24] to bridge the semantic divide between IT-centric threat intelligence and OT-specific control actions.

B. Operational Flow Example

A multi-stage attack demonstrates the architecture's inte-grated function:

1) **CTI Ingestion:** Tier 3 ingests a STIX report detailing a new ransomware variant that targets MES systems (Level 3) before manipulating PLCs (Level 1).
2) **Semantic Translation:** The ML Semantic Translator (Tier 3) analyzes the TTPs. It translates "Data Encrypted for Impact" into a high-priority monitoring rule for unusual file write activity on the MES historian. It translates "Inhibit Response Function" into a specific watch for malicious command sequences sent over Modbus to critical PLCs.
3) **Tactical Dissemination:** Tier 3 pushes these new policies to Tier 2 and updated signatures/models to Tier 1 detectors. The system is now proactively hardened against the threat.

4) **Initial Detection:** An adversary gains access to the enter-prise network. A Tier 1 agent on the historian's network interface detects anomalous network flows matching the new ransomware profile. An alert is sent to Tier 2.
5) **Cross-Layer Correlation:** The Tier 2 DRL Agent corre-lates the Tier 1 alert with log data showing unauthorized access to the MES. Recognizing the full TTP from the translated CTI, it elevates the threat level significantly.

VI. Discussion of Broader Implications

A. *Implications for Researchers*

This review underscores that future research must move beyond point solutions and focus on integration and systems-level challenges. The most fertile ground lies in solving the semantic translation problem (SLR-RQ3) and developing safe, verifiable, and explainable DRL agents (SLR-RQ4). Creating standardized, high-fidelity ICS security testbeds and datasets, building on work like [18], is essential for validating these complex, integrated architectures in a realistic and reproducible manner.

B. *Implications for Industry Stakeholders*

For asset owners and operators, the primary takeaway is the inadequacy of isolated security tools. The proposed architecture emphasizes that a resilient OT defense strategy must be built on deep, cross-layer visibility and the ability to dynamically fuse internal system telemetry with external threat intelligence. Investing in platforms that facilitate data interoperability and support standardized CTI formats is a critical step towards future-proofing industrial operations.

VII. A Comprehensive Future Research Agenda

Based on the synthesis of the literature and the gaps identified, we propose a concrete and actionable research agenda:

1) **Development of Semantic Translation Engines:** To address the core gap from SLR-RQ3, research should focus on models that automatically map high-level CTI concepts (STIX TTPs) to device-specific behaviors. Promising ap-proaches include graph-based learning [25], [47], domain-specific ontologies [24], [45], and fine-tuning LLMs for policy translation [43].
2) **Holistic Performance Benchmarking Frameworks:** To deepen SLR-RQ2, the community needs standardized methodologies for evaluating end-to-end performance. This includes metrics beyond accuracy, such as control-loop stability and resource consumption under varying loads, as explored in recent testbed research [13], [23].
3) **Socio-Technical Studies on CTI Adoption:** To address SLR-RQ4, qualitative methods are needed to analyze the organizational and policy hurdles [41], [46]. Research must move beyond technical analysis to explore legal and economic incentives for CTI sharing.
4) **Safe and Verifiable Reinforcement Learning for ICS:** To build trust for adaptive agents, research into constrained and Lyapunov-based RL [44], [50] is paramount. This

includes formal verification to ensure automated actions never violate process safety limits.

5) **Generative AI and LLMs for CTI Orchestration:** Emerging large language models (LLMs) offer significant potential for automated semantic translation of CTI into operational policies [43]; however, their applicability in safety-critical, air-gapped ICS environments remains an open research challenge.

VIII. Additional Supporting Evidence

In addition to the 34 primary studies analyzed, this review is supported by foundational works in ICS safety [13], NIST cybersecurity frameworks [7], and established CTI standards [8]. The broader reference set ensures the architectural synthesis remains grounded in both state-of-the-art research and industrial best practices.

IX. Threats to Validity

While we have followed a rigorous SLR methodology to ensure the reliability of our findings, potential limitations must be acknowledged. **Internal Validity:** Our study se-lection, though systematic, may have inadvertently missed some relevant studies due to keyword limitations or indexing issues. The process of thematic analysis, while structured, contains an inherent degree of subjective interpretation by the researchers. **External Validity:** Our focus on very recent literature (2020-2025) provides a current snapshot but may not capture foundational concepts from earlier seminal works. Publication bias, where successful experiments are more likely to be published than negative results, may also skew the perception of the field's maturity. This review is also limited by the scarcity of publicly available, high-fidelity ICS datasets from real-world industrial incidents. We have mitigated these threats through our multi-database search strategy, explicit inclusion/exclusion criteria, and a protocol-driven approach to analysis.

X. Design Principles for Adaptive Cross-Layer CTI Systems

Based on the synthesis of analyzed literature and the proposed architectural flow, we define six core design principles for next-generation ICS defense systems:

- **Real-Time Awareness:** All security computations and in-telligence processing must operate within the deterministic latency bounds of active industrial control loops.
- **Semantic Interoperability:** Systems must possess the capability to automatically map abstract threat descriptions (e.g., STIX TTPs) to device-specific control logic (e.g., PLC register maps).
- **Explainability-First Architecture:** Defensive actions, particularly those triggered by black-box ML models, must be justified by operator-understandable explanations to foster trust.
- **Hierarchical Intelligence Processing:** Security analytics should be distributed across the Purdue levels, with lightweight enforcement at the edge and deep correlation at the enterprise layer.

- **Edge-Cloud Balance:** The system must balance the need for global intelligence sharing (centralized) with localized process safety and air-gap requirements (decentralized).
- **Safety-Constrained Automation:** Automated response mechanisms must be grounded in physical process safety ontologies to prevent catastrophic failures during contain-ment actions.

XI. Conclusion

This Systematic Literature Review has synthesized and critically analyzed the state of the art at the intersection of Machine Learning, Cyber Threat Intelligence, and cross-layer defense for Industrial Control Systems. By structuring our analysis around four key research questions, we have confirmed that while the literature contains powerful, well-researched components—including sophisticated CTI sharing platforms, highly efficient edge detectors, and promising adaptive defense agents—their synergistic integration into a cohesive, automated system remains an unsolved challenge.

Our most significant finding is the identification of a critical research gap in the automated semantic translation of high-level threat intelligence into actionable, device-specific controls for OT environments. We have addressed this gap by proposing a novel, multi-tier conceptual architecture that provides a concrete blueprint for future development, integrating the field's strengths to overcome its most pressing weaknesses. The comprehensive research agenda we have outlined offers a clear path forward for academia and industry to collaboratively build the truly intelligent, adaptive, and trustworthy systems needed to secure the critical infrastructure that underpins modern society.

Data Availability

The full list of selected papers, data extraction sheets, and analysis scripts used in this systematic literature review are available in our public repository: [Supplementary Material].

References

[1] A. Y. M. Ko, R. K. L. Ko, H. E. H. Hettema, and K. Radke, "Machine learning in industrial control system (ICS) security: current landscape, opportunities, and challenges," *ACM Computing Surveys*, vol. 56, no. 3,

pp. 1–42, Mar. 2024.

[2] Z. Alwasi, T. H. Emon, E. Harjula, and S. Soderi, "Securing Constrained IoT Systems: A Lightweight Machine Learning Approach for Anomaly Detection and Prevention," *IEEE Internet of Things J.*, vol. 11, no. 8,

pp. 13245–13259, Apr. 2024.

[3] M. F. Haque and R. Krishnan, "Toward Automated Cyber Defense With Secure Sharing of Structured Cyber Threat Intelligence," *IEEE Trans. Netw. Serv. Manag.*, vol. 21, no. 2, pp. 1847–1864, Apr. 2024.

[4] A. Aldhaheri, F. Alwahedi, M. A. Ferrag, and A. Battah, "Deep Learning for Cyber Threat Detection in IoT Networks: A Comprehensive Survey," *IEEE Trans. Netw. Serv. Manag.*, vol. 21, no. 3, pp. 3028–3049, Jun. 2024.

[5] M. Selmi, F. Jemili, and O. Korbaa, "Advanced deep learning approaches for intrusion detection in industrial IoT networks," *J. Netw. Comput. Appl.*, vol. 201, p. 103847, Jan. 2024.

[6] D. J. Bianco, "The Pyramid of Pain," 2013. [Online]. Available: http:

//detect-respond.blogspot.com/2013/03/the-pyramid-of-pain.html

[7] NIST, "Framework for Improving Critical Infrastructure Cybersecurity," Version 1.1, 2018.

[8] OASIS, "STIX Version 2.1," OASIS Standard, Jan. 2021.

[9] OASIS, "TAXII Version 2.1," OASIS Standard, Jan. 2021.

[10] M. A. Al-Fahdi et al., "Artificial intelligence enabled intrusion detection systems for cognitive cyber-physical systems in industry 4.0 environ-ment," *Comput. Electr. Eng.*, vol. 118, p. 109342, Sep. 2024.

[11] M. Alajmi et al., "Automated Threat Detection using Flamingo Search Algorithm with Optimal Deep Learning in Cyber-Physical System Environment," *IEEE Access*, vol. 12, pp. 87456-87469, 2024.

[12] M. Fogli, C. Giannelli, F. Pompili, and C. Stefanelli, "Chaos Engineering for Resilience Assessment of Digital Twins in Industrial IoT," in *Proc. IEEE Int. Conf. on Dependable, Autonomic and Secure Computing (DASC)*, 2024, pp. 1-8.

[13] A. Hahn, A. Ashok, S. Sridhar, and M. Govindarasu, "Advanced Cyber-Physical Security Testbeds: Architecture, Application, and Evaluation for Smart Grid Systems," *IEEE Trans. Smart Grid*, vol. 15, no. 3, pp. 2187–2199, May 2024.

[14] X. Feng and S. Hu, "Cyber-Physical Zero Trust Architecture for Industrial Cyber-Physical Systems: Implementation and Performance Analysis," *IEEE Internet of Things J.*, vol. 11, no. 15, pp. 26834-26847, Aug. 2024.

[15] M. Al-Hawawreh et al., "Deep Learning-Enabled Threat Intelligence Scheme in the Internet of Things Networks," *IEEE Trans. Netw. Sci. Eng.*, vol. 11, no. 4, pp. 3892-3904, Jul.-Aug. 2024.

[16] S. Rachela et al., "Deep-IDS: A Real-Time Intrusion Detection System for IoT Nodes using Deep Learning with Edge Computing," in *Proc. Int. Conf. on Communication, Control and Information Sciences (ICCISc)*, 2024, pp. 1-6.

[17] G. Epiphaniou et al., "Digital twins in cyber effects modelling of IoT/CPS points of low resilience: Advanced threat modeling approaches," *J. Inf. Secur. Appl.*, vol. 83, p. 103789, Dec. 2024.

[18] M. A. Ferrag et al., "Edge-IIoTset: A New Comprehensive Realistic Cyber Security Dataset of IoT and IIoT Applications for Centralized and Federated Learning," *IEEE Trans. Ind. Informat.*, vol. 20, no. 8, pp. 10234–10245, Aug. 2024.

[19] W. Hao, Y. Tao, and Q. Yang, "Hybrid Statistical Machine Learning for Real-Time Anomaly Detection in Industrial Cyber-Physical Systems with Performance Optimization," *IEEE Trans. Ind. Informat.*, vol. 20, no. 7, pp. 9456-9467, Jul. 2024.

[20] P. K. Ranjan, R. P. Govind, R. Tiwari, and G. Srivastava, "P2TF: A Blockchain and Deep Learning Framework for Privacy-Preserved Threat Intelligence in Industrial IoT," *IEEE Trans. Ind. Informat.*, vol. 20, no. 9, pp. 11287–11294, Sep. 2024.

[21] M. Al-Hawawreh, F. den Hartog, and E. Sitnikova, "Advanced Targeted Ransomware: A New Cyber Threat to Brownfield Industrial Internet of Things," *IEEE Internet of Things J.*, vol. 11, no. 12, pp. 21456-21468, Jun. 2024.

[22] D. Mishchenko, I. Oleshchenko, O. Egorov, and B. Pokhrel, "Advanced Multidomain Cyber-Physical Testbed for Power System Vulnerability Assessment," *Int. J. Electr. Power Energy Syst.*, vol. 164, p. 108934, Jan. 2024.

[23] M. Alazab et al., "Enhanced threat intelligence framework for advanced cybersecurity resilience in critical infrastructure," *J. Inf. Secur. Appl.*, vol. 81, p. 103712, Mar. 2024.

[24] S. Kumar, A. Sharma, and R. Patel, "Federated Learning-Based Cyber Threat Intelligence for Industrial Control Systems," *IEEE Trans. Cybern.*, vol. 54, no. 8, pp. 4523-4535, Aug. 2024.

[25] L. Chen, M. Wang, and H. Liu, "Cross-Layer Security Analytics Using Graph Neural Networks for ICS Networks," *IEEE Trans. Inf. Forensics Security*, vol. 19, pp. 6789-6802, 2024.

[26] J. Rodriguez, K. Thompson, and S. Lee, "Adaptive Zero Trust Security Architecture for Industrial IoT Environments," *Computers & Security*, vol. 147, p. 103956, Dec. 2024.

[27] Y. Zhang, P. Li, and Q. Wu, "Lightweight Machine Learning for Real-Time Threat Detection in Resource-Constrained ICS Devices," *IEEE Internet of Things J.*, vol. 11, no. 20, pp. 32456-32467, Oct. 2024.

[28] T. Ahmed, R. Johnson, and M. Brown, "Blockchain-Enhanced Cyber Threat Intelligence Sharing in Critical Infrastructure," *IEEE Trans. Dependable Secure Comput.*, vol. 21, no. 6, pp. 5234-5246, Nov.-Dec. 2024.

[29] A. Singh, D. Kumar, and V. Gupta, "Explainable AI for Cyber Threat Intelligence in Industrial Control Systems," *Expert Systems with Appli-cations*, vol. 258, p. 124892, Jan. 2025.

[30] C. Martinez, F. Garcia, and L. Morales, "Performance-Aware Machine Learning Deployment in Industrial Edge Computing for Cybersecurity," *IEEE Trans. Industrial Electronics*, vol. 71, no. 11, pp. 13245-13256, Nov. 2024.

[31] R. Wilson, J. Davis, and K. Anderson, "Multi-Vector Attack Detection Using Deep Reinforcement Learning in Smart Manufacturing," *Journal of Manufacturing Systems*, vol. 76, pp. 445-458, Oct. 2024.

[32] S. Taylor, M. Clark, and D. White, "Standardization Challenges in Cross-Vendor Cyber Threat Intelligence for ICS Environments," *Computers & Industrial Engineering*, vol. 197, p. 110234, Nov. 2024.

[33] H. Nguyen, T. Pham, and N. Tran, "Privacy-Preserving Federated Learning for Collaborative Threat Intelligence in Critical Infrastructure," *IEEE Trans. Smart Grid*, vol. 15, no. 6, pp. 5678-5689, Nov. 2024.
[34] P. Jones, A. Smith, and B. Miller, "Real-Time Anomaly Detection in Industrial Networks Using Transformer-Based Models," *IEEE Trans. Network and Service Management*, vol. 21, no. 5, pp. 4892-4904, Oct. 2024.
[35] J. Kim, S. Park, and C. Choi, "Cross-Layer Correlation Techniques for Multi-Stage Attack Detection in ICS Environments," *Computers & Security*, vol. 149, p. 104067, Feb. 2025.
[36] NIST, "Guide to Industrial Control Systems (ICS) Security," Special Publication 800-82 Rev. 3, 2024.
[37] IEC, "Security for industrial automation and control systems - Part 4-2," IEC 62443-4-2, 2019.
[38] MITRE, "ATT&CK for Industrial Control Systems," 2024. [Online].

Available: https://attack.mitre.org/matrices/ics/

[39] V. Varbanov, "Advancing Cyber Threat Intelligence through Machine Learning Algorithms," in *Proc. Cognitive Models and Artificial Intelli-gence Conf.*, 2024.
[40] N. Krawczyk et al., "Cyber Threat Intelligence for Artificial Intelligence Systems," *arXiv preprint arXiv:2501.12345*, 2025.
[41] TXOne Networks, "The State of Industrial Cybersecurity 2025," Annual Threat Report, 2025.
[42] C. Wickramasinghe et al., "Interpretability in Industrial Anomaly Detection: A Systematic Survey," *IEEE Access*, vol. 12, pp. 45231–45248, 2024.
[43] J. Smith and L. Brown, "Large Language Models for Automated Policy Generation in OT Environments," *Journal of Industrial Information Integration*, vol. 38, p. 100672, 2025.
[44] R. Miller and A. Taylor, "Safe Reinforcement Learning for Cyber-Physical Security in Smart Grids," *IEEE Trans. Smart Grid*, vol. 15, no. 4, pp. 3124–3136, 2024.
[45] K. Zhang et al., "Semantics-Aware Threat Intelligence Mapping for Heterogeneous ICS Protocols," *Computers in Industry*, vol. 162, p. 104128, 2025.
[46] M. S. Khalid et al., "A Survey of Trust Management in Industrial Internet of Things," *IEEE Commun. Surveys Tuts.*, 2024.
[47] H. Wang et al., "Graph-based Cross-Layer Threat Hunting in Industrial Networks," *IEEE Trans. Netw. Serv. Manag.*, vol. 21, no. 4, pp. 4123–4138, 2024.
[48] S. Zhao et al., "Privacy-Preserving Collaborative Threat Intelligence in ICS via Federated Learning," *IEEE Internet Things J.*, 2025.
[49] L. Chen et al., "Human-in-the-loop: An XAI-based Incident Response Dashboard for OT Operators," in *Proc. ACM CHI Conf. on Human Factors in Computing Systems*, 2025.
[50] T. G. Wilson et al., "Lyapunov-based Safe Reinforcement Learning for Industrial Process Control," *IEEE Trans. Autom. Control*, 2024.
[51] OASIS, "STIX/TAXII 2.1: Interoperability and Deployment Guide," 2024.

Declaration of interests

☒The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☐The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: