Membrane Technology
ISSN (online): 1873-4049

Zero-Trust and Artificial Intelligence-Driven Security Strategies for Cyber-Physical Systems in Pharmaceutical and Defense Facilities

Prasanth Alluri

Principle IT Security Architect,California, USA

Abstract

Cyber-physical systems form the operational backbone of pharmaceutical manufacturing and defense facilities, where tightly coupled digital and physical processes enable precision, efficiency, and mission-critical performance. However, the convergence of information technology and operational technology has expanded the attack surface, exposing these environments to sophisticated cyber threats capable of causing physical disruption, safety hazards, regulatory noncompliance, and mission failure. Traditional perimeter-based security models are increasingly inadequate for protecting such complex and distributed systems. This article examines the application of Zero-Trust Architecture combined with artificial intelligence driven security strategies as a unified approach for strengthening cyber-physical system protection in high assurance environments. The proposed approach integrates continuous identity and device verification, micro- segmentation, and least-privilege access with AI-based behavioral analytics, anomaly detection, and dynamic risk scoring. By aligning enforcement oriented zero-trust controls with data-driven intelligence, the framework enables early threat detection, limits lateral movement, and supports rapid, risk-aware response while respecting operational safety constraints. The article synthesizes existing standards and research, develops an integrated architectural model tailored to pharmaceutical and defense contexts, and analyzes sector- specific implementation considerations. Key contributions include a structured mapping of zero-trust principles to cyber-physical system layers, the role of AI in operational security monitoring, and a comparative assessment of security priorities across regulated industrial domains. The findings highlight how coordinated Zero-Trust and AI strategies can enhance resilience, visibility, and trustworthiness in cyber-physical systems without compromising safety or operational continuity.

Keywords: *Zero Trust Architecture; cyber-physical systems; operational technology security; artificial intelligence driven cybersecurity; anomaly detection; pharmaceutical manufacturing security; defense facility security; industrial control systems standards.*

1. Introduction

1.1 Background: Cyber-Physical Systems in high-risk facilities (OT and IT convergence)

Cyber-Physical Systems (CPS) integrate computational elements, communication networks, and physical processes to enable real-time monitoring and control of industrial operations. In sectors such as pharmaceuticals and defense, CPS form the backbone of mission-critical activities, including automated manufacturing lines, quality control systems, secure logistics, weapons support infrastructure, and facility management systems. Traditionally, these environments relied on isolated Operational Technology (OT) networks designed primarily for reliability, safety, and deterministic performance rather than cybersecurity.

Over the past two decades, however, increasing digitalization and the demand for operational efficiency have driven the convergence of OT with Information Technology (IT) systems. Enterprise resource planning platforms, manufacturing execution systems, remote maintenance tools, and cloud-based analytics are now routinely connected to industrial control systems, programmable logic controllers, and sensor networks.

While this convergence improves visibility, scalability, and productivity, it also significantly expands the attack surface of CPS environments (Humayed et al., 2017; Stouffer et al., 2023).

1.2 Why pharmaceutical and defense facilities are uniquely high-stakes

Pharmaceutical and defense facilities represent two of the most sensitive CPS application domains due to the critical nature of their outputs and the severity of potential failures. In pharmaceutical manufacturing, CPS govern batch production, environmental controls, quality assurance processes, and data integrity mechanisms that directly affect patient safety and regulatory compliance. A cybersecurity incident in this context can compromise drug quality, disrupt supply chains, invalidate clinical or manufacturing records, and lead to significant public health risks (U.S. Food and Drug Administration, 2023).

Defense facilities face an equally high, and often higher, level of risk. CPS in defense environments support mission-critical operations, secure communications, logistics, weapons systems maintenance, and infrastructure protection. Cyber intrusions targeting these systems may aim to degrade operational readiness, exfiltrate sensitive information, or cause physical disruption with national security implications. Defense CPS are also frequent targets of advanced persistent threats, insider attacks, and sophisticated state-sponsored adversaries, making conventional perimeter-based security models insufficient (U.S. Department of Defense, 2022).

1.3 Motivation for Zero Trust and AI-driven security in CPS

Traditional CPS security architectures have relied heavily on network segmentation and assumed trust within defined perimeters. However, the erosion of clear network boundaries due to remote access, third-party vendors, cloud connectivity, and mobile workforces has rendered implicit trust models increasingly ineffective. Zero Trust Architecture (ZTA) challenges this assumption by enforcing continuous verification of identities, devices, and sessions, regardless of network location (Rose et al., 2020; Syed et al., 2022).

While ZTA provides a strong policy and enforcement foundation, its effectiveness in CPS environments depends on timely and accurate situational awareness. This requirement has driven growing interest in Artificial Intelligence (AI) and machine learning techniques for CPS security. AI-driven approaches enable the analysis of large volumes of heterogeneous CPS data, including network traffic, control commands, and sensor telemetry, to identify anomalous behaviors and emerging threats that may evade signature-based detection (Vinayakumar et al., 2019; Buczak C Guven, 2016).

1.4 Research aim, scope, and key contributions

The primary aim of this research is to examine how Zero Trust Architecture and AI-

driven security strategies can be jointly applied to enhance the protection of Cyber-Physical Systems in pharmaceutical and defense facilities. The study adopts a conceptual and applied perspective, focusing on architectural design, operational considerations, and sector-specific requirements rather than proposing a single implementation or algorithm.

The scope of the paper includes CPS environments characterized by OT and IT convergence, with particular attention to industrial control systems, sensor networks, and supervisory platforms commonly used in pharmaceutical manufacturing and defense infrastructure. The analysis draws on established standards, federal guidance, and peer-reviewed research to ensure relevance and practical applicability.

1.5 Paper organization

The remainder of this paper is organized as follows. Section 2 presents the conceptual foundations of CPS security, Zero Trust Architecture, and AI-driven cybersecurity. Section 3 analyzes the threat landscape and attack pathways relevant to pharmaceutical and defense CPS environments. Section 4 discusses the design of Zero Trust architectures adapted to CPS constraints, while Section 5 examines AI-driven security strategies suitable for operational technology contexts. Section 6 introduces an integrated Zero Trust and AI security framework, followed by sector-specific implementation discussions in Sections 7 and 8. Section 9 provides a comparative analysis of security priorities across pharmaceutical and defense facilities. Sections 10 through 12 address evaluation metrics, implementation challenges, and practical deployment recommendations. Finally, Section 13 outlines future research directions, and Section 14 concludes the paper.

2. Conceptual Foundations

This section establishes the theoretical and architectural basis for securing cyber-physical systems in pharmaceutical and defense facilities by integrating established CPS security models with Zero Trust Architecture and artificial intelligence–driven security mechanisms. Together, these foundations support a shift from perimeter-based protection to continuous, risk-aware enforcement aligned with modern threat realities.

2.1 Cyber-Physical Systems Security Model

Cyber-physical systems integrate computational elements with physical processes through sensors, actuators, controllers, and communication networks. In pharmaceutical manufacturing plants and defense facilities, CPS environments typically include programmable logic controllers, distributed control systems, supervisory control and data acquisition platforms, manufacturing execution systems, and enterprise IT systems. The tight coupling between digital control and physical outcomes means that cybersecurity incidents can directly result in safety hazards, product integrity failures, or mission disruption.

A widely adopted conceptual model for structuring CPS and industrial control system security is the Purdue Enterprise Reference Architecture. The Purdue model organizes systems into hierarchical levels, beginning with physical processes and field devices at the lowest levels and extending upward to enterprise and business systems. This layered view supports security zoning by clearly separating operational technology from information technology domains.

2.2 Zero Trust Architecture Essentials

Zero Trust Architecture is a security paradigm that eliminates implicit trust and requires continuous verification of all access requests, regardless of network location. Instead of assuming that assets inside a defined perimeter are trustworthy, ZTA enforces access decisions based on identity, context, and risk at every interaction.

Identity is the central control point in ZTA. In CPS environments, identity extends beyond human users to include devices, applications, and automated processes. Controllers, sensors, remote maintenance tools, and data historians must all be uniquely identifiable and authenticated before being permitted to communicate or execute commands.

Device trust complements identity by assessing the security posture and integrity of endpoints. In operational environments, this may include validating firmware versions, configuration states, and communication behavior of industrial devices. Devices that fail posture checks or exhibit abnormal behavior can be restricted or isolated without disrupting the entire system.

2.3 AI-Driven Security Essentials

Artificial intelligence provides the analytical capability needed to operationalize Zero Trust in complex CPS environments. Unlike conventional signature-based tools, AI-driven security systems learn normal patterns of behavior and identify deviations that may indicate malicious activity or system failure.

Anomaly detection is a core AI application in CPS security. By modeling baseline behavior of network traffic, control commands, and sensor readings, AI systems can detect subtle deviations that traditional rule-based approaches may overlook. This is especially valuable in operational networks where attacks often masquerade as legitimate control activity.

Behavioral analytics extend anomaly detection by correlating activities across users, devices, and processes. For example, unusual sequences of PLC commands, unexpected timing of batch control operations, or atypical remote access behavior can be flagged for investigation. Behavioral models are well suited to environments where deterministic processes dominate, such as pharmaceutical batch manufacturing or defense control systems.

2.4 Threat Actor Models Relevant to CPS

Threat modeling for cyber-physical systems must account for a diverse set of adversaries with varying capabilities and objectives. External attackers include cybercriminal groups and advanced persistent threat actors targeting intellectual property, sabotage opportunities, or strategic disruption. These actors often exploit remote access services, vulnerable edge devices, or trusted third-party connections to gain initial access.

Insider threats pose a particularly serious risk in CPS environments. Authorized personnel, contractors, or operators may intentionally or inadvertently misuse access privileges. Because insiders often operate within trusted zones, their actions can bypass traditional perimeter defenses and directly impact physical processes.

Table 1. ZTA principles mapped to CPS security layers

CPS layer	Key assets	ZTA control objective	Practical controls	Standards reference
Field devices	Sensors, actuators, PLCs, embedded controllers	Establish device identity and integrity	Device authentication, firmware validation, secure boot, network allowlisting	NIST SP 800-82; IEC 62443-4-2
Control layer	PLC networks, DCS controllers, safety systems	Enforce least privilege and limit lateral movement	Micro-segmentation, role- based access, command whitelisting	IEC 62443-3-3; ISA 62443-3-2
Supervisory layer	SCADA servers, historians, HMI systems	Continuous verification of users and applications	Strong authentication, session monitoring, behavior analytics	NIST SP 800-53 Rev. 5; NIST SP 800-207
Enterprise layer	MES, ERP, identity services, SOC tools	Context-aware access and risk-based enforcement	Identity federation, policy engines, AI-driven risk scoring	NIST CSF 2.0; CISA Zero Trust Maturity Model

Mapping Zero Trust control objectives to cyber-physical system layers with practical controls and relevant standards.

3. Threat Landscape and Attack Pathways in Pharmaceutical and

Defense CPS

Cyber-Physical Systems deployed in pharmaceutical manufacturing and defense facilities operate at the intersection of digital control and physical processes. The increasing convergence of information technology and operational technology has significantly expanded the attack surface of these environments. Unlike traditional enterprise IT systems, CPS environments are constrained by real-time requirements, safety considerations, and long equipment lifecycles, making conventional security controls difficult to apply uniformly. As a result, adversaries increasingly exploit weaknesses that allow them to traverse from cyber domains into physical process control, producing impacts that extend beyond data compromise to operational disruption and safety hazards.

3.1 Common CPS and OT Threat Vectors

Several recurring threat vectors dominate cyber incidents in CPS and OT environments across both pharmaceutical and defense sectors.

- **Remote access misuse** remains one of the most prevalent entry points. Remote connectivity is often required for vendors, maintenance personnel, and system integrators to support distributed operations. However, weak authentication mechanisms, shared credentials, lack of session monitoring, and excessive privileges frequently enable attackers to hijack legitimate access channels. Once compromised, these connections provide a trusted pathway into environments that are otherwise segmented from the public internet.

- **Lateral movement** is a critical escalation technique following initial access. Many CPS networks rely on flat or weakly segmented architectures designed for operational efficiency rather than security. Attackers exploit trust relationships between systems to move laterally from enterprise IT assets to OT networks, often using standard administrative tools or compromised credentials. This movement enables reconnaissance of control systems and identification of high-value targets such as programmable logic controllers, supervisory control systems, and historian databases.

3.2 Facility-Specific Risks

While common threat vectors apply across CPS deployments, the consequences and risk profiles differ significantly between pharmaceutical and defense facilities due to differences in operational objectives, regulatory environments, and threat actors.

In **pharmaceutical facilities**, CPS security failures directly threaten product integrity and patient safety. Manufacturing processes rely on tightly controlled batch operations governed by validated control systems and electronic records. Unauthorized manipulation of process parameters can compromise batch integrity, leading to substandard or contaminated products. Even subtle data integrity attacks, such as

altering sensor readings or batch logs, can undermine quality systems without immediately triggering alarms. Additionally, prolonged production downtime caused by ransomware or control system disruption can lead to drug shortages, regulatory noncompliance, and significant financial losses. The combination of strict regulatory oversight and complex supply chains further increases the impact of cyber incidents in this sector.

3.3 Intrusion Chain Logic for CPS Environments

Attacks against CPS environments typically follow a multi-stage intrusion chain that begins in traditional IT domains and progressively advances toward physical process manipulation. Understanding this chain is essential for designing effective Zero Trust and AI-driven security strategies.

The intrusion process often starts with **initial access** through phishing, compromised remote access credentials, or exploitation of internet-facing services in the enterprise network. Once inside, attackers conduct **internal reconnaissance** to identify network topology, trust relationships, and pathways leading toward OT environments. This phase frequently involves credential harvesting and privilege escalation.

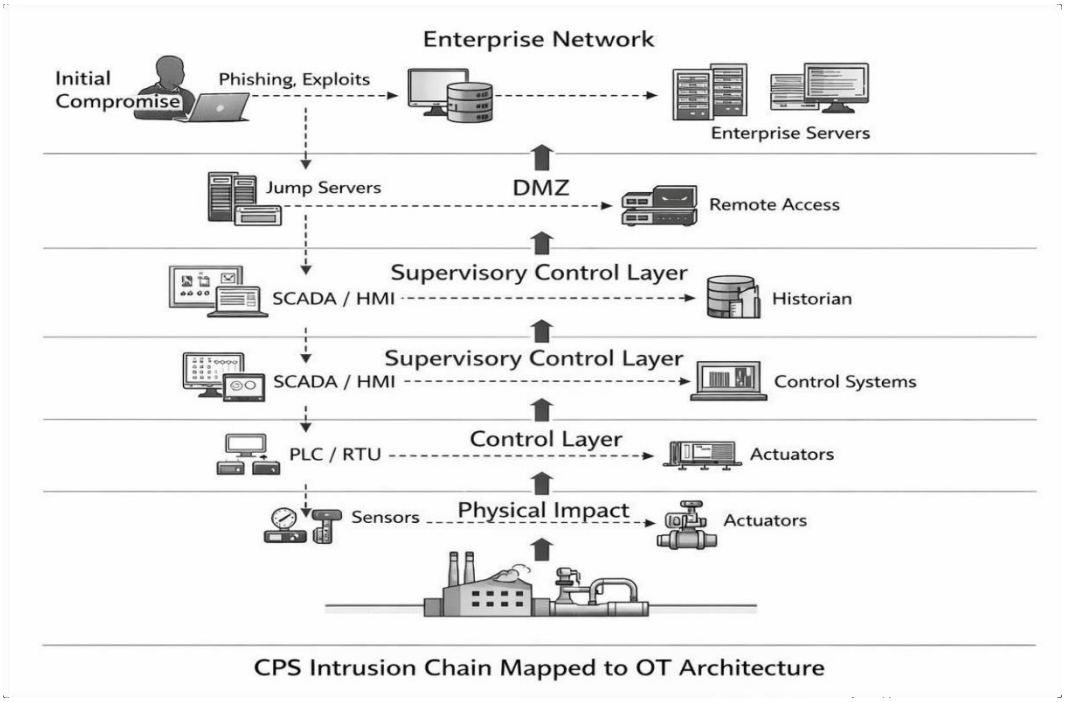

Figure 1. CPS Intrusion Chain Mapped to OT Architecture

This diagram illustrates the typical progression of a cyber intrusion in cyber-physical systems, showing how attackers move from enterprise IT environments through the DMZ into operational technology layers. It highlights the transition from initial compromise and lateral movement to control system manipulation and eventual physical process impact, emphasizing the critical security boundary crossings within pharmaceutical and defense CPS environments.

4. Zero-Trust Architecture Design for CPS Environments

Zero Trust Architecture (ZTA) represents a fundamental shift from perimeter-based security toward a model in which no user, device, or workload is implicitly trusted. While ZTA principles are well established in enterprise IT environments, their application to cyber- physical systems (CPS) and operational technology (OT) environments requires careful adaptation. Pharmaceutical and defense facilities rely on safety-critical, deterministic, and often legacy-controlled systems, making direct transplantation of IT-centric Zero Trust models impractical. This section presents a tailored Zero Trust design approach that accounts for OT constraints while preserving the core principles of continuous verification, least privilege, and explicit trust decisions.

4.1 Adapting Zero Trust Architecture to OT Constraints

Operational technology environments differ fundamentally from conventional IT systems in terms of design priorities and operational requirements. CPS environments emphasize **safety, availability, and determinism**, often above confidentiality. Consequently, Zero Trust mechanisms must be introduced in a manner that does not disrupt time-sensitive control processes or compromise physical safety.

One of the primary constraints is **latency sensitivity**. Industrial control loops operate on millisecond-level timing requirements, and additional authentication or inspection steps introduced by ZTA controls can introduce unacceptable delays. Therefore, enforcement points in OT networks must be positioned strategically, typically at zone boundaries rather than inline with real-time control traffic. Policy decisions are often precomputed and cached to minimize runtime overhead.

4.2 Identity-Centric Access Control for Human and Machine Identities

Identity is the foundation of Zero Trust, and in CPS environments this extends beyond human users to include machines, applications, and autonomous components. Effective ZTA design requires consistent identification, authentication, and authorization across all interacting entities.

For **human identities**, role-based and attribute-based access control mechanisms are applied, reflecting job function, operational context, and risk level. In pharmaceutical facilities, this includes operators, quality assurance personnel, maintenance engineers, and external auditors. In defense facilities, access policies must also consider clearance levels, mission roles, and operational states. Strong authentication mechanisms such as multi- factor authentication are typically enforced at higher network layers, such as engineering workstations and remote access gateways, rather than directly on control devices.

4.3 Micro-Segmentation and Policy Enforcement Points

Micro-segmentation is a core Zero Trust strategy that limits lateral movement by isolating workloads and devices into tightly controlled security zones. In CPS environments, segmentation aligns naturally with established industrial models such as the Purdue Enterprise Reference Architecture.

Rather than broad network segments, Zero Trust encourages **fine-grained segmentation** based on function, criticality, and risk exposure. For example, in pharmaceutical manufacturing, batch control systems, environmental monitoring systems, and quality data repositories are segmented into distinct zones with strictly defined communication paths.

In defense facilities, mission-critical systems are isolated from administrative and support networks, even if they reside within the same physical infrastructure.

Policy enforcement points are deployed at key boundaries, including:

- Between IT and OT networks
- Between OT supervisory and control layers
- At remote access ingress points
- At application interfaces and data repositories

4.4 Continuous Verification and Device Posture Assessment

Continuous verification is essential to Zero Trust, particularly in environments where threats may evolve over long operational lifetimes. In CPS contexts, verification extends beyond user authentication to include **ongoing assessment of device and system posture**.

Device posture assessment evaluates whether a component remains in a trusted state by monitoring factors such as firmware integrity, configuration changes, communication patterns, and compliance with approved baselines. In OT environments, where traditional endpoint detection agents are often unsuitable, posture assessment is frequently performed through passive monitoring and protocol analysis.

Behavioral consistency plays a central role. Because CPS systems exhibit predictable communication patterns, deviations from established baselines may indicate misconfiguration, unauthorized access, or malicious activity. Continuous verification allows trust decisions to be adjusted dynamically, enabling rapid containment without relying solely on predefined static rules.

4.5 Secure Remote Access Model for Vendors and Contractors

Remote access represents one of the highest-risk entry points in CPS environments, especially in pharmaceutical and defense facilities that rely on external vendors for equipment maintenance, calibration, and software updates. Zero Trust

mandates that remote access be treated as inherently untrusted, regardless of network location.

A secure Zero Trust remote access model incorporates:

- ❖ Strong identity verification for vendors and contractors
- ❖ Explicit session authorization tied to approved tasks
- ❖ Time-limited and purpose-specific access
- ❖ Continuous session monitoring and recording

Remote access is typically mediated through hardened gateways that enforce policy decisions before allowing interaction with OT assets. Direct connections to control devices are avoided. Instead, access is brokered through controlled jump hosts or secure application proxies that restrict commands and data flows.

Table 2. Zero-Trust control categories and implementation in OT environments

ZTA pillar	OT-specific challenge	Recommended control	Deployment note	Operational risk
Identity	Limited authentication support in legacy PLCs	Centralized identity services with gateway- based enforcement	Apply identity checks at network or application gateways	Misattributed trust if device identity is inferred incorrectly
Devices	Inability to install endpoint agents	Passive monitoring and firmware integrity checks	Use protocol-aware sensors and asset inventories	Delayed detection of compromised devices
Network	Deterministic traffic and latency sensitivity	Micro-segmentation with static allow lists	Enforce at zone boundaries rather than inline control paths	Misconfigured policies may disrupt operations
Applications	Tight coupling between control software and hardware	Application-level access proxies and command validation	Restrict access to approved functions only	Limited flexibility during emergency operations
Data	Real-time process data cannot tolerate inspection delays	Policy-based data access and secure replication	Apply controls at historian and data aggregation layers	Exposure if aggregation points are compromised

5. AI-Driven Security Strategies for CPS Protection

Artificial intelligence provides a critical analytical layer for securing cyber-physical systems by enabling continuous monitoring, adaptive detection, and context-aware response across tightly coupled digital and physical processes. In pharmaceutical and defense facilities, where CPS environments generate large volumes of heterogeneous operational data and tolerate minimal disruption, AI-driven security strategies must be carefully designed to align with safety, reliability, and compliance constraints. This section examines the data foundations for AI in CPS, the detection models best suited to these environments, response automation mechanisms, and the risks posed by adversarial machine learning.

5.1 Data Sources for AI in CPS

Effective AI-driven security in CPS depends on the availability and quality of operational data that accurately reflects both cyber activity and physical process behavior. Unlike traditional IT systems, CPS environments rely on multiple specialized data sources that operate at different time scales and levels of determinism.

❖ **Network flow data** captures communication patterns between CPS components, including programmable logic controllers, supervisory control systems, human- machine interfaces, and enterprise gateways. These flows provide visibility into command-and-control channels, lateral movement attempts, and deviations from expected communication paths. In CPS environments, network flows are particularly valuable because normal traffic patterns are often highly regular and protocol- specific, making anomalous behavior easier to identify when appropriate baselines are established.

❖ **Historian logs** store time-series records of process variables, alarms, and system states over extended periods. These logs are essential for understanding long-term operational trends, correlating cyber events with physical outcomes, and training AI models on realistic process behavior. In pharmaceutical manufacturing, historian data reflects batch execution parameters and quality-relevant variables, while in defense facilities it may capture mission system states and environmental conditions.

5.2 Detection Models Suited to CPS

Detection models for CPS security must account for the structured, repetitive, and safety- critical nature of industrial and mission systems. Traditional signature-based approaches are insufficient in these environments, as novel attacks and subtle process manipulations often lack predefined indicators. As a result, anomaly-focused AI models are particularly well suited to CPS protection.

❖ **Unsupervised anomaly detection** is widely applicable in CPS contexts because labeled attack data is scarce and difficult to obtain without disrupting operations. These models learn normal system behavior directly from historical data and flag deviations that exceed acceptable thresholds. Techniques

such as clustering, density estimation, and reconstruction-based models are commonly used to identify abnormal network traffic patterns, unexpected control commands, or unusual sensor readings. In CPS environments, unsupervised methods are effective when normal operations are stable and well defined, but they require careful tuning to avoid excessive false positives during legitimate process changes.

❖ **Semi-supervised behavior baselines** combine limited labeled data with large volumes of normal operational data. These models establish explicit baselines for acceptable behavior, such as typical command sequences or process variable ranges, and then evaluate new observations against these learned profiles. Semi- supervised approaches are particularly valuable in pharmaceutical and defense facilities where certain abnormal conditions, such as maintenance activities or controlled shutdowns, are known and can be incorporated into training data. By incorporating domain knowledge, these models achieve higher precision than purely unsupervised methods while remaining adaptable to evolving operations.

In practice, CPS security deployments often use a combination of unsupervised and semi- supervised models. Unsupervised models provide broad coverage for unknown threats, while semi-supervised baselines refine detection accuracy for critical assets and workflows. This layered approach supports early detection without overwhelming operators with low- confidence alerts.

5.3 Response Automation and Human-in-the-Loop Controls

Detection alone is insufficient in CPS environments, where delayed or inappropriate responses can have physical safety or mission consequences. AI-driven security strategies therefore integrate response mechanisms that balance automation with human oversight.

❖ **Response automation** enables rapid containment actions such as session termination, network segmentation enforcement, or access privilege reduction when high-confidence anomalies are detected. In Zero Trust-aligned CPS architectures, AI outputs can dynamically inform policy enforcement points, restricting access or isolating affected components in near real time. Automation is most appropriate for well-defined, low-risk actions that have been pre-approved through safety and operational reviews.

❖ **Human-in-the-loop controls** remain essential for decisions that may impact physical processes or critical missions. Security analysts, OT engineers, or mission operators review AI-generated alerts, validate contextual information, and authorize escalated responses such as system shutdowns or control logic changes. Human oversight ensures that automated actions do not inadvertently disrupt legitimate operations or violate regulatory requirements.

5.4 Adversarial ML Threats in Operational Environments

While AI enhances CPS security, it also introduces new attack surfaces in the form of adversarial machine learning threats. These risks are particularly concerning in operational environments where AI models directly influence security decisions.

❖ **Poisoning attacks** target the training data used to build or update AI models. In CPS environments, attackers may inject manipulated sensor readings or altered logs to skew behavior baselines, causing malicious activity to appear normal. Poisoning is especially dangerous in systems that rely on continuous or online learning, as gradual data manipulation can remain undetected over long periods.

❖ **Evasion attacks** aim to craft malicious actions that intentionally mimic normal behavior to bypass detection models. In CPS contexts, this may involve carefully timed command sequences or subtle physical parameter adjustments that remain within learned thresholds. Evasion attacks exploit the predictability of AI models and highlight the need for diverse detection features and periodic model evaluation.

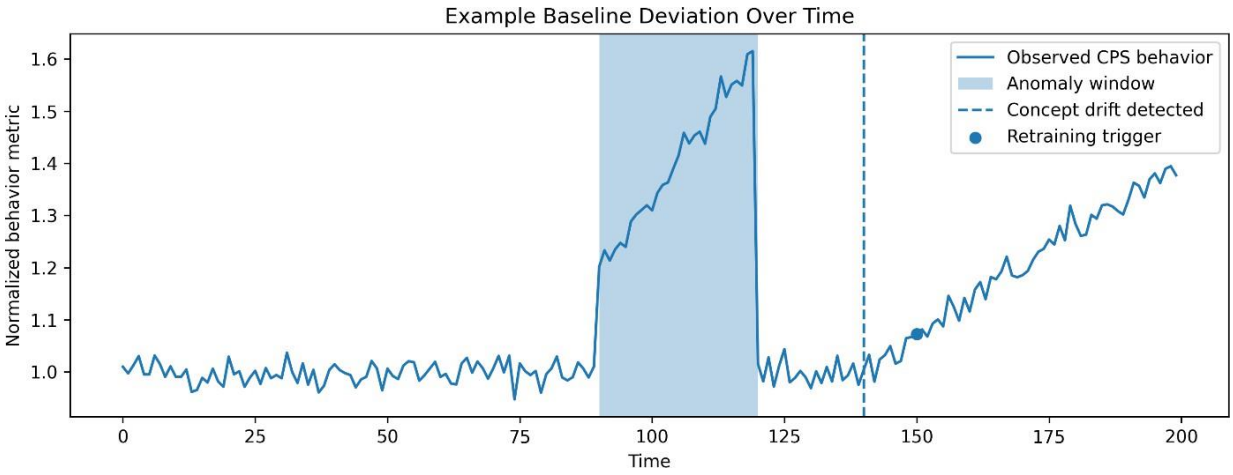

Figure 2. Example baseline deviation over time

Illustration of a learned CPS behavioral baseline, highlighting an anomaly window with significant deviation, a detected concept drift point, and a retraining trigger used to maintain AI model reliability in operational environments.

6. Integrated Framework: Zero Trust and AI-Driven Security for High-Assurance Facilities

High-assurance environments such as pharmaceutical manufacturing plants and defense facilities require security architectures that can simultaneously enforce strict access controls and adapt to evolving cyber-physical threats. Traditional perimeter-based defenses are insufficient in these contexts due to the convergence of IT and OT, legacy system constraints, and the high consequences of both cyber and physical failures. This section proposes an **integrated Zero Trust and AI-driven security framework** designed specifically for cyber-physical systems, combining deterministic policy enforcement with adaptive, intelligence-driven threat

detection.

6.1 Unified Architecture for Enforcement and Intelligence

The proposed framework integrates **Zero Trust Architecture (ZTA)** enforcement mechanisms with **artificial intelligence-based security analytics** into a unified control loop. Rather than treating access control and threat detection as separate functions, the architecture tightly couples them through a continuous feedback process.

At the core of the framework are three interacting components:

❖ **Telemetry Collection Layer**

This layer aggregates multi-source data from CPS environments, including:

- Network traffic between IT, DMZ, and OT zones
- Controller commands and sensor telemetry from PLCs and field devices
- Authentication and authorization logs from identity systems
- Application and historian logs from manufacturing execution systems and supervisory platforms

6.2 Trust Evaluation Logic

A defining feature of the integrated framework is its **multi-dimensional trust evaluation model**, which moves beyond simple identity verification. Trust is continuously assessed across four complementary dimensions.

❖ **Identity Trust:** Identity trust evaluates the legitimacy of human and non-human identities, including operators, engineers, service accounts, and automated processes. It incorporates factors such as authentication strength, role alignment, historical behavior patterns, and compliance with least-privilege principles.

❖ **Device Trust:** Device trust assesses the security posture of CPS assets, including controllers, sensors, gateways, and engineering workstations. Attributes include firmware integrity, configuration state, communication patterns, and deviation from known operational baselines. This is particularly important in environments with long-lived legacy devices.

❖ **Session Trust:** Session trust evaluates the real-time context of interactions, such as access timing, command frequency, protocol usage, and network path. Even a trusted identity on a trusted device may be denied or restricted if session behavior deviates from expected operational norms.

6.3 Dynamic Risk Scoring Model

To support adaptive decision-making, the framework employs a **dynamic risk scoring model** that translates trust evaluations and anomaly signals into actionable

security decisions. The model incorporates four primary factors.

- **Impact:** Impact represents the potential physical, safety, operational, or mission consequences if a given activity were malicious. For example, unauthorized changes to a pharmaceutical batch control system or a defense logistics controller carry significantly higher impact than anomalies in peripheral monitoring systems.

- **Asset Criticality:** Asset criticality reflects the role of a device, application, or process within the CPS environment. Assets directly affecting safety, product quality, or mission assurance are weighted more heavily than supporting infrastructure components.

- **Confidence:** Confidence measures the reliability of the detection signal. It accounts for model accuracy, signal consistency across data sources, and historical false- positive rates. This reduces unnecessary disruption caused by uncertain alerts in sensitive OT environments.

6.4 SOC Integration Workflow in High-Noise Environments

Security Operations Centers supporting CPS environments face uniquely high alert volumes and limited tolerance for false positives. The proposed framework is designed to integrate seamlessly into SOC workflows while reducing cognitive and operational burden.

AI-driven analytics prioritize alerts based on composite risk scores rather than raw anomaly counts. Low-risk deviations are logged for trend analysis, while high-risk events trigger actionable alerts enriched with contextual information such as affected assets, likely objectives, and recommended responses.

The Zero Trust policy engine enables **graduated response strategies**, allowing SOC analysts to enforce constraints incrementally rather than resorting to disruptive shutdowns. Human-

in-the-loop controls are maintained for high-impact decisions, particularly those affecting safety-critical operations.

By aligning AI intelligence with ZTA enforcement, the SOC shifts from reactive alert handling to **risk-driven operational decision-making**, which is essential in high-assurance pharmaceutical and defense environments.

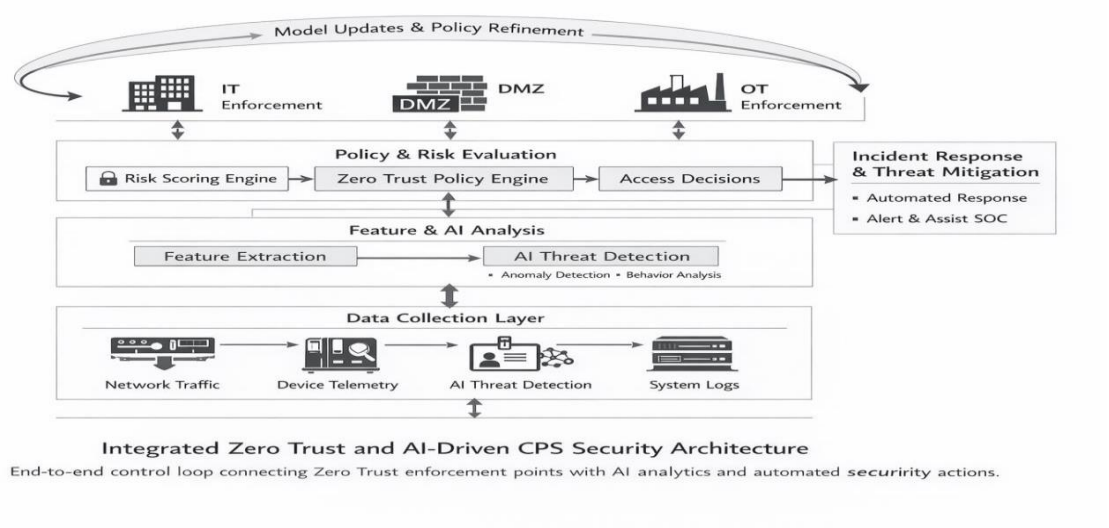

Figure 3. Integrated Zero Trust and AI-Driven CPS Security Architecture

This diagram illustrates an integrated Zero Trust and AI-driven security architecture tailored for cyber-physical systems in high-assurance environments. Continuous telemetry from network traffic, devices, identities, and system logs is analyzed using AI-based feature extraction and threat detection, with results fed into a Zero Trust policy and risk evaluation engine. Security decisions are enforced across IT, DMZ, and OT zones, while incident response outcomes loop back to refine policies and retrain models, enabling adaptive, risk- aware protection without disrupting critical operations.

Table 3. AI Methods Mapped to Security Functions in Cyber-Physical Systems

AI approach	Detection target	CPS data input	Strength	Limitation	Deployment note
Unsupervised anomaly detection	Unknown or novel attacks	Network flows, sensor telemetry, command sequences	Detects zero-day behaviors without labeled data	May produce false positives during process changes	Requires stable baseline periods
Semi-supervised learning	Known attack patterns with variation	Labeled incident data, operational logs	Balances accuracy and adaptability	Limited by quality of labeled data	Effective for SOC- validated datasets
Time-series modeling	Process manipulation and drift	Sensor readings, control loop variables	Captures temporal dependencies	Sensitive to noise and missing data	Best suited for continuous processes

Graph-based analytics	Lateral movement and dependency abuse	Network topology, asset relationships	Reveals complex attack paths	Computationally intensive at scale	Useful for blast-radius estimation
Ensemble models	Multi-vector attacks	Combined CPS telemetry	Improves robustness and confidence	Increased complexity	Recommended for high-criticality assets

7. Implementation in Pharmaceutical Facilities

7.1 Cyber-Physical System Components in Pharmaceutical Manufacturing

Pharmaceutical manufacturing environments rely on tightly integrated **cyber-physical systems (CPS)** to ensure product quality, safety, and regulatory compliance. These CPS environments typically consist of a layered architecture combining operational technology and information systems that collectively support batch production, continuous manufacturing, and quality assurance activities.

At the operational layer, **programmable logic controllers (PLCs)** execute deterministic control logic governing mixers, reactors, filling lines, and environmental controls. PLCs interface directly with **sensors and actuators** that measure temperature, pressure, flow rate, humidity, and chemical composition, all of which are critical quality attributes in pharmaceutical production. These field-level components are often legacy devices with limited native security capabilities, making them attractive targets for cyber intrusion (Stouffer et al., 2011; Stouffer et al., 2023).

Above the control layer, **supervisory control and data acquisition (SCADA)** systems provide centralized monitoring, alarm management, and operator interaction. SCADA platforms aggregate data from PLCs and enable supervisory control actions, such as recipe changes or process parameter adjustments. In parallel, **Manufacturing Execution Systems (MES)** coordinate production scheduling, enforce batch recipes, manage electronic batch records, and synchronize production with enterprise resource planning systems. MES platforms are especially critical because they serve as the digital bridge between validated production processes and business operations.

7.2 Applying Zero Trust Architecture Controls to Pharmaceutical Workflows

Applying **Zero Trust Architecture (ZTA)** to pharmaceutical CPS requires adapting identity- centric security principles to environments where availability, determinism, and validation are paramount. Unlike traditional perimeter-based approaches, ZTA

assumes no implicit trust and enforces continuous verification across users, devices, and applications (Rose et al., 2020; Syed et al., 2022).

Role-based access control is foundational in pharmaceutical workflows. Operators, quality assurance personnel, maintenance engineers, and external vendors must be granted only the minimum privileges required for their tasks. ZTA enforces this principle by binding access decisions to strong identity verification, contextual attributes, and device posture rather than static network location. For example, an engineer performing maintenance on a filling line PLC may be granted time-limited access restricted to specific functions and devices, with all actions logged for auditability.

7.3 Artificial Intelligence Use Cases in Pharmaceutical CPS Security

Artificial intelligence plays a complementary role to ZTA by providing **continuous, data- driven insight** into system behavior that static rules alone cannot capture. In pharmaceutical CPS environments, AI models are particularly well suited to detecting subtle deviations that may indicate cyber manipulation or process tampering.

One primary use case is **batch anomaly detection**. By learning normal production patterns from historical sensor data and batch records, AI models can identify deviations in process timing, parameter relationships, or sequence execution that fall outside expected variability.

7.4 Compliance and Quality System Alignment Considerations

Any security strategy deployed in pharmaceutical facilities must align with established **quality systems and regulatory expectations**. Cybersecurity controls should be implemented in a manner that preserves data integrity, system availability, and traceability, which are central to regulatory oversight. Guidance from regulatory authorities emphasizes the importance of cybersecurity risk management as part of the overall quality system lifecycle, particularly for computerized systems that influence product quality (U.S. Food and Drug Administration, 2023).

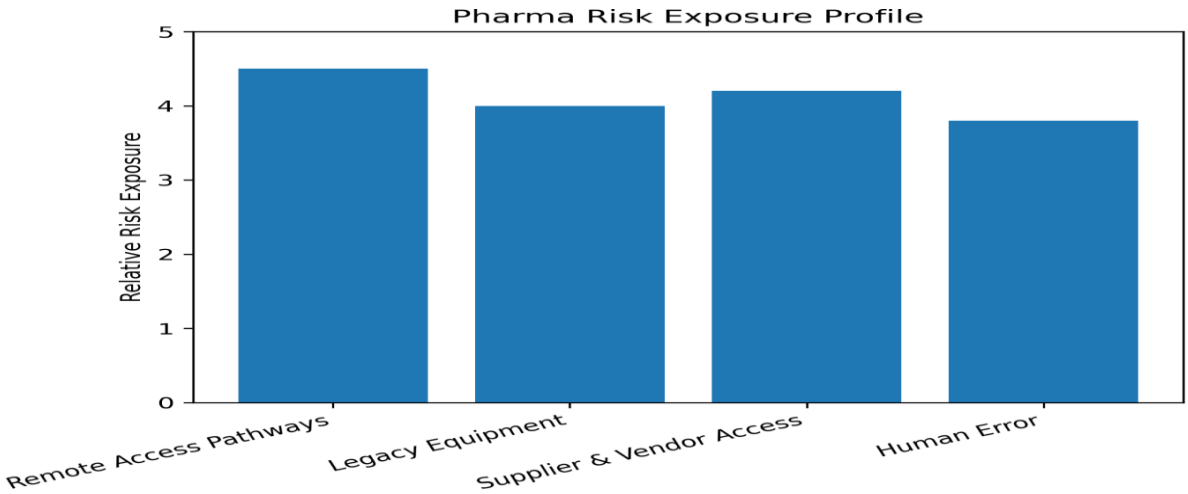

Figure 4. Pharma Risk Exposure Profile

8. Implementation in Defense Facilities

8.1 Defense Cyber-Physical System Context

Defense facilities operate CPS environments designed to support **mission-critical functions** where confidentiality, integrity, and availability are tightly coupled to national security objectives. These systems often control physical assets such as radar installations, weapons systems, logistics infrastructure, and energy or communications platforms. Unlike commercial environments, defense CPS are frequently deployed within **restricted or classified networks**, with stringent assurance requirements and limited tolerance for failure (U.S. Department of Defense, 2022).

8.2 Applying Zero Trust Architecture Pillars in Defense Environments

Zero Trust Architecture aligns closely with defense security principles by enforcing **strong identity assurance**, continuous verification, and strict segmentation across all system layers. In defense CPS, identity extends beyond human users to include devices, applications, and workloads, each of which must be uniquely identifiable and continuously authenticated (Rose et al., 2020).

- **Strong identity enforcement** typically involves multi-factor authentication, cryptographic credentials, and hardware-backed identity for devices and endpoints. Access decisions are dynamically evaluated based on mission context, threat intelligence, and system state rather than static network boundaries.

- **Network and workload segmentation** is used to compartmentalize mission systems, reducing the blast radius of any compromise. Even within classified environments, ZTA enforces explicit trust relationships, ensuring that compromise of one subsystem does not automatically grant access to adjacent assets.

8.3 Artificial Intelligence Use Cases in Defense CPS Security

Artificial intelligence enhances defense CPS security by enabling **proactive threat detection and adaptive response** in complex, high-noise environments. One prominent use case is **AI-driven threat hunting**, where models analyze large volumes of telemetry to surface weak signals associated with advanced persistent threats. These techniques support analysts by prioritizing high-confidence indicators rather than relying solely on predefined signatures.

AI also supports **deception-aware detection**, correlating adversary interactions with decoy systems or misleading signals to infer intent and capability. By identifying how attackers probe or avoid certain assets, AI systems can inform defensive posture adjustments and improve attribution.

8.4 Operational Continuity and Incident Containment Model

Defense CPS security strategies prioritize **operational continuity under attack**. ZTA enforcement points and AI analytics are integrated into incident response

workflows that emphasize containment, graceful degradation, and rapid recovery. When anomalous activity is detected, automated policies can restrict access, isolate affected components, and alert operators while preserving essential mission functions.

Incident containment strategies are supported by predefined playbooks and continuous monitoring, ensuring that responses are both timely and aligned with operational priorities. This approach reflects a shift from purely preventive security toward resilient system design capable of operating in contested environments (Joint Task Force, 2020).

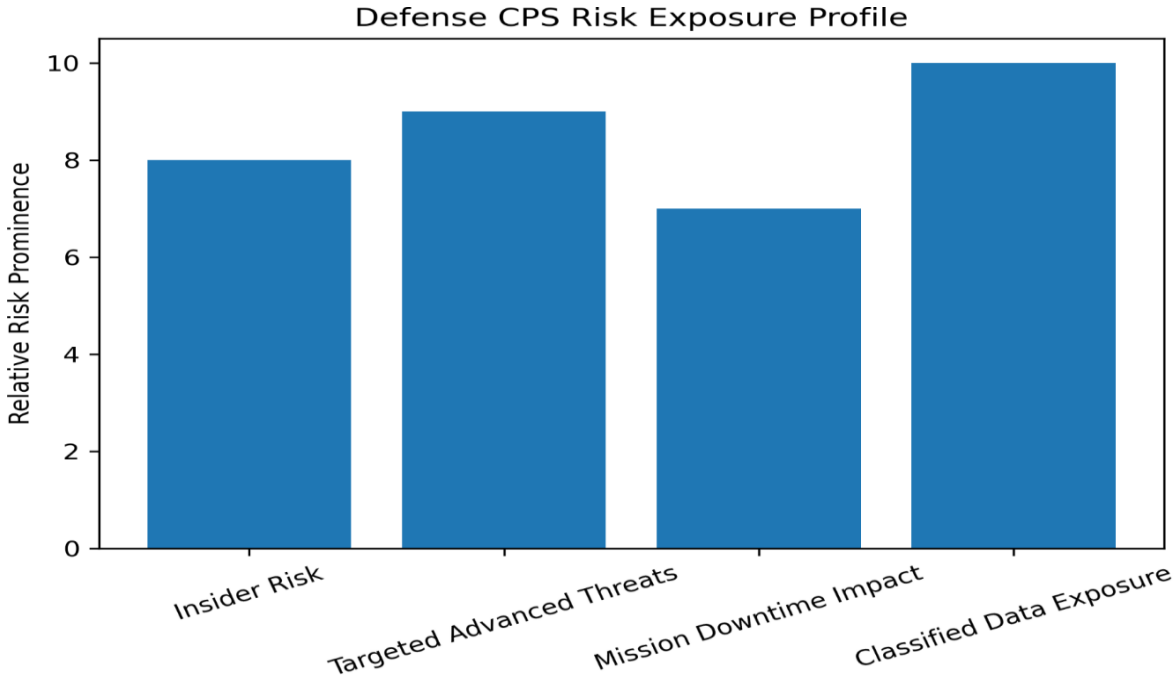

Figure 5 . Defense Risk Exposure Profile

G. Comparative Analysis: Pharmaceutical vs Defense CPS Security Priorities

Cyber-physical systems deployed in pharmaceutical manufacturing and defense facilities operate under distinct operational missions, regulatory regimes, and threat models. While both sectors are classified as high-consequence environments, the **objectives of attackers, acceptable operational risk, and definitions of security success differ in important ways**, which directly influences how Zero-Trust Architecture (ZTA) and AI-driven security mechanisms should be designed and evaluated.

G.1 Similarities and Differences in Threat Objectives

At a fundamental level, **both pharmaceutical and defense CPS environments face advanced, persistent, and well-resourced adversaries**. Common threat

objectives across both sectors include unauthorized access to operational networks, lateral movement across segmented systems, manipulation of control logic, and disruption of cyber-physical processes. Insider threats, compromised third-party access, and exploitation of legacy OT components are also shared risk vectors.

However, **the ultimate objectives of attackers diverge** between the two sectors. In pharmaceutical facilities, threat actors are often motivated by **economic gain, intellectual property theft, production sabotage, or regulatory disruption**. Attacks targeting batch records, manufacturing execution systems, or environmental control systems can undermine product quality, patient safety, and regulatory compliance, even if the immediate physical impact appears limited.

9.2 Differences in Governance, Tolerance for Downtime, and Incident Response

Governance structures significantly shape CPS security strategies. Pharmaceutical facilities operate under **strict regulatory oversight**, including quality management systems, validation requirements, and auditability mandates. Changes to systems, including security controls, often require formal validation and documentation. Consequently, **incident response procedures must balance rapid containment with regulatory traceability and product impact assessments**.

Defense facilities, by contrast, function under **centralized command-and-control governance models** with classified operational constraints. While compliance requirements are stringent, they are typically aligned with national security frameworks rather than commercial regulatory bodies. This allows for **more rapid, authority-driven response actions**, including aggressive network isolation or system shutdowns when mission integrity is at risk.

Tolerance for downtime also differs. Pharmaceutical production environments may tolerate short, controlled outages if necessary to protect product integrity or prevent contamination, provided regulatory reporting requirements are met. Defense systems, however, often operate under **near-zero tolerance for downtime during active missions**, requiring security controls that emphasize graceful degradation and fail-safe operation rather than full shutdown.

G.3 Sector

The concept of cybersecurity success in CPS environments is inherently **context-dependent.**

For pharmaceutical facilities, success is defined by:

- **Integrity** of production data and control logic
- **Safety** of manufactured products and environments
- **Compliance** with regulatory and quality standards
- **Continuity** of validated manufacturing operations

Security mechanisms are considered effective when they prevent unauthorized changes, detect deviations from validated process states, and support auditability without disrupting approved workflows.

In defense facilities, success is defined by:

- **Mission assurance**, ensuring operational objectives can be achieved despite cyber threats
- **Confidentiality**, protecting classified data and operational intent
- **Resilience**, maintaining functional capability under attack conditions

Here, security success is measured less by regulatory alignment and more by the system's ability to **absorb, adapt to, and recover from hostile cyber activity without compromising mission outcomes.**

These differing success criteria underscore the need for sector-aware ZTA policies and AI evaluation metrics, rather than one-size-fits-all security benchmarks.

Table 4. Sector Comparison Matrix: Pharmaceutical vs Defense CPS Security

Dimension	Pharmaceutical Facilities	Defense Facilities	Implication for ZTA	Implica D Sc
Primary threat objective	IP theft, production sabotage, regulatory disruption	Mission degradation, intelligence gathering, strategic impact	Strong identity and access controls with auditability	Emphasis anomaly d
Governance model	Regulatory and quality- driven	Command-and-control, security-driven	Policy enforcement aligned with compliance workflows	AI decisio support t
Downtime tolerance	Limited but manageable with controls	Extremely low during missions	Segmentation and isolation must be selective	AI respons false shut
Incident response style	Structured, documented, compliance-oriented	Rapid, authority-driven, mission-focused	ZTA enforcement points must support escalation paths	AI prioritiz high- conf
Definition of success	Integrity, safety, compliance, continuity	Mission assurance, confidentiality, resilience	Policies emphasize least privilege and validation	AI focuses threat de

10. Evaluation and Metrics

Evaluating the effectiveness of Zero-Trust and AI-driven security in CPS environments requires **metrics that reflect both technical performance and operational impact**. Traditional IT security metrics are insufficient on their own, as CPS security failures can have physical, safety, and mission-level consequences.

10.1 What to Measure for Zero-Trust Success

Key indicators of effective ZTA implementation in CPS environments include:

- **Access policy compliance**, measuring how consistently access requests align with defined identity, device, and contextual policies
- **Segmentation effectiveness**, assessing whether unauthorized lateral movement is prevented across CPS zones and conduits
- **Least privilege coverage**, evaluating the proportion of users, devices, and services operating under minimal required access

These metrics collectively indicate whether trust is being **continuously evaluated and enforced**, rather than assumed based on network location or static credentials.

10.2 What to Measure for AI Security Success

AI-driven security mechanisms must be evaluated using metrics that balance detection capability with operational feasibility:

- **Precision and recall**, to quantify detection accuracy and completeness
- **False positive rates**, particularly critical in OT environments where unnecessary alerts increase operator fatigue
- **Mean Time to Detect (MTTD)** and **Mean Time to Respond (MTTR)** improvements, reflecting real operational gains

In CPS contexts, **high precision is often more valuable than maximal recall**, as excessive false alarms can erode trust in automated systems and lead to alert suppression.

10.3 Practical Validation Approach

Because real-world CPS attacks are rare and high-risk to reproduce, **validation relies on controlled and simulated methods**, including:

- **Testbeds**, such as representative OT or CPS environments that emulate real process behavior
- **Red-team exercises**, simulating adversarial behavior across IT-OT boundaries
- **Tabletop simulations**, enabling cross-disciplinary evaluation of decision-making, escalation, and response coordination

These approaches allow organizations to evaluate both **technical controls and**

human- machine interaction under realistic conditions.

10.4 Operational Monitoring and Model Drift Management

AI models deployed in CPS environments must account for **concept drift**, as operational behavior evolves due to maintenance, upgrades, or process optimization. Continuous monitoring should therefore include:

- ❖ Performance tracking against baseline behavior
- ❖ Periodic retraining triggers tied to verified operational changes
- ❖ Human review loops for model updates affecting safety-critical decisions

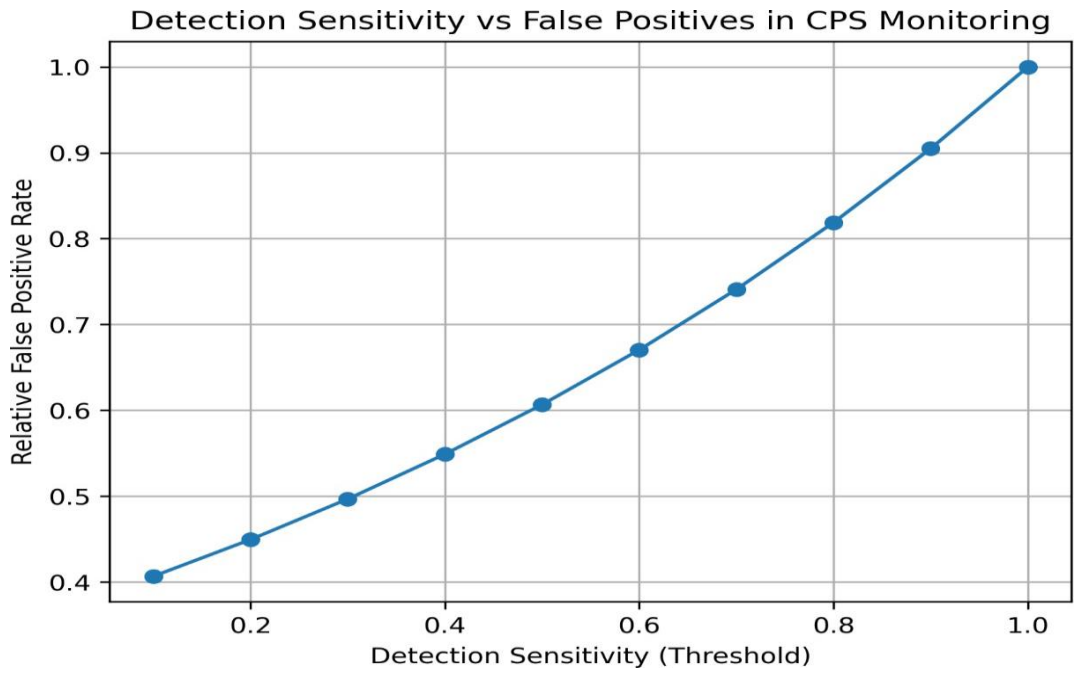

Figure 6. Detection Sensitivity vs False Positives in CPS Monitoring

This figure illustrates the expected trade-off in CPS environments: as detection sensitivity increases, the false positive rate rises, highlighting why threshold calibration is critical for SOC sustainability and operator trust.

11. Challenges, Limitations, and Risk Management

Despite their promise, Zero-Trust and AI-driven security strategies face **practical limitations**

in real CPS deployments.

- ❖ **Legacy Systems and Downtime Constraints:** Many CPS environments rely on legacy controllers and protocols that lack native support for modern security controls. Retrofitting ZTA enforcement or telemetry collection must be done cautiously to avoid unintended downtime.

- **Data Quality, Labeling, and Incomplete Visibility:** AI effectiveness depends on data quality. In CPS environments, incomplete logging, proprietary protocols, and limited labeling hinder model accuracy and generalization.

- **Safety vs Security Tradeoffs in OT:** Security controls that interrupt or delay control signals may introduce safety risks. Any automated response must therefore be **safety-aware and context-sensitive**.

- **AI Risks: Explainability, Adversarial Attacks, and Drift:** Opaque models can reduce operator trust, while adversarial manipulation and model drift can degrade performance over time. Explainability and validation mechanisms are essential in safety-critical domains.

12. Recommendations and Practical Roadmap

The effective deployment of Zero-Trust Architecture combined with AI-driven security analytics in cyber-physical systems requires a structured, phased approach. Pharmaceutical and defense facilities operate under strict safety, regulatory, and availability constraints, which makes incremental implementation essential. This section presents a practical roadmap that balances security enhancement with operational continuity, followed by a governance model and a discussion of minimum viable versus mature security states.

12.1 Step-by-Step Deployment Roadmap

- **Phase 1: Asset Inventory and Segmentation:** The foundation of Zero-Trust in CPS environments is complete visibility. Organizations must first establish a comprehensive inventory of assets across IT and OT domains, including PLCs, sensors, actuators, servers, workstations, and remote access gateways. Network segmentation should then be applied to separate enterprise IT systems from operational networks and to further isolate critical control zones. In pharmaceutical and defense contexts, this phase reduces the attack surface and limits the blast radius of potential intrusions while preserving deterministic control traffic.

- **Phase 2: Identity and Access Hardening:** Once assets are identified and segmented, identity becomes the primary control plane. Human users, devices, applications, and services should be assigned strong, verifiable identities with role- based and context-aware access policies. Least-privilege access must be enforced for operators, engineers, vendors, and automated processes. In regulated pharmaceutical environments, this phase directly supports auditability and change control, while in defense facilities it strengthens insider threat mitigation and access traceability.

12.2 Governance Model

Successful implementation depends on clear governance and cross-functional

collaboration. Security responsibilities in CPS environments span multiple organizational domains, each with distinct priorities.

- **Security Operations Center (SOC):** Monitors alerts, investigates anomalies, and coordinates incident response across IT and OT environments.
- **OT Engineers:** Ensure that security controls do not compromise safety, availability, or real-time process requirements.
- **Quality Assurance:** In pharmaceutical facilities, validates that security changes comply with quality systems, documentation, and regulatory expectations.
- **Compliance and Risk Management:** Aligns technical controls with standards, policies, and regulatory obligations.
- **Leadership:** Provides strategic direction, risk acceptance decisions, and resource allocation.

A shared governance framework ensures that Zero-Trust and AI controls are deployed consistently while respecting operational realities.

12.3 Minimum Viable Controls Versus Mature State Controls

Not all organizations can immediately achieve a fully mature Zero-Trust and AI-driven security posture. A minimum viable state focuses on essential protections such as basic segmentation, strong authentication for remote access, and limited anomaly detection. A mature state extends these controls through pervasive identity enforcement, fine-grained micro-segmentation, advanced AI analytics, and automated response mechanisms.

Recognizing this progression allows organizations to prioritize investments and demonstrate incremental risk reduction.

Table 5: Practical Deployment Roadmap for Zero-Trust and AI-Driven CPS Security

Phase	Objective	Key Actions	Required Stakeholders	Expected O
Phase 1	Establish visibility and containment	Asset discovery, network mapping, zone segmentation	OT engineers, IT security, operations	Reduced a and limited movement
Phase 2	Enforce least-privilege access	Identity assignment, role-based access, strong authentication	SOC, IAM teams, compliance	Controlled auditable a assets

Phase 3	Enable adaptive detection	Telemetry collection, AI baseline modeling, anomaly detection	SOC, data science, OT engineers	Early detec and proces
Phase 4	Sustain resilience	Model retraining, policy refinement, security exercises	SOC, leadership, risk management	Continuous improveme term resili

13. Future Research Directions

Despite recent advances, several research challenges remain at the intersection of Zero- Trust, AI, and cyber-physical systems. First, **cross-domain threat intelligence for CPS** requires further exploration to enable information sharing between IT, OT, and sector- specific ecosystems without compromising confidentiality. Second, **secure federated learning in OT environments** offers promise for collaborative model training across facilities while preserving data sovereignty. Third, **explainable AI for safety-critical anomaly detection** is essential to build trust among engineers and regulators who must understand and validate automated decisions. Finally, **automated policy generation and verification for Zero-Trust architectures** represents an important direction for reducing configuration errors and ensuring consistent enforcement at scale.

14. Conclusion

This study has presented a structured approach to securing cyber-physical systems in pharmaceutical and defense facilities through the integration of Zero-Trust Architecture and AI-driven security strategies. The proposed roadmap demonstrates how identity-centric controls, continuous verification, and adaptive analytics can be applied incrementally without disrupting critical operations.

Key takeaways include the necessity of asset visibility and segmentation as a foundation, the central role of identity in CPS security, and the value of AI in detecting subtle behavioral deviations that traditional controls may miss. While pharmaceutical and defense facilities differ in regulatory and mission priorities, both benefit from a unified, risk-based security model that emphasizes resilience and accountability.

The closing recommendation is that organizations adopt Zero-Trust and AI-driven security not as isolated technologies, but as an integrated operational philosophy, supported by governance, continuous learning, and sector-specific risk awareness.

References

1. National Institute of Standards and Technology. (2023). Artificial Intelligence Risk Management Framework (AI RMF 1.0) (NIST AI 100-1). https://nvlpubs. nist. gov/nistpubs/ai/nist. ai, 100-1.
2. Alexander, O., Belisle, M., C Steele, J. (2020). MITRE ATTCCK for industrial control

systems: Design and philosophy. The MITRE Corporation: Bedford, MA, USA, 29, 21-85. https://attack.mitre.org/docs/ATTACK_for_ICS_Philosophy_March_2020.pdf

3. Boyens, J., Paulsen, C., Moorthy, R., Bartol, N., C Shankles, S. A. (2015). Supply chain risk management practices for federal information systems and organizations. NIST Special publication, 800(161), 32.
4. Buczak, A. L., C Guven, E. (2015). A survey of data mining and machine learning methods for cyber security intrusion detection. IEEE Communications surveys C tutorials, 18(2), 1153-1176.
5. Carlini, N., C Wagner, D. (2017, May). Towards evaluating the robustness of neural networks. In 2017 ieee symposium on security and privacy (sp) (pp. 39-57). Ieee.
6. Cichonski, P., Millar, T., Grance, T., C Scarfone, K. (2012). Computer security incident handling guide. NIST Special Publication, 800(61), 1-147.
7. CIO Council. (2024). Federal Zero Trust Data Security Guide. U.S. Federal CIO Council. https://www.cio.gov/assets/files/Zero-Trust-Data-Security-Guide_Oct24-Final.pdf
8. Prasanth Alluri. (2022, December). Data-Driven and Artificial Intelligence-Enabled Frameworks for Sustainable Energy, Rural Transportation Networks, and Water Resource Management in Developing Economies. (IJCNIS) : https://www.ijcnis.org/index.php/ijcnis/article/view/8807
9. CISA. (2025). Zero Trust Architecture Implementation (Federal guidance). U.S. Department of Homeland Security. https://www.dhs.gov/sites/default/files/2025-04/2025_0129_cisa_zero_trust_architecture_implementation.pdf
10. Cybersecurity and Infrastructure Security Agency. (2023). Zero Trust Maturity Model (Version 2.0). U.S. Department of Homeland Security.https://www.cisa.gov/sites/default/files/2023-04/CISA_Zero_Trust_Maturity_Model_Version_2_508c.pdf
11. Dempsey, K. L., Chawla, N. S., Johnson, L. A., Johnston, R., Jones, A. C., Orebaugh, A. D., ... C Stine, K. M. (2011). Sp 800-137. information security continuous monitoring (iscm) for federal information systems and organizations.
12. Force, J. T. (2018). Risk management framework for information systems and organizations. NIST Special Publication, 800, 37.
13. Force, J. T. (2020). Control baselines for information systems and organizations. NIST Special Publication, 800, 53B.
14. Force, J. T. (2020). Security and privacy controls for information systems and organizations (No. NIST Special Publication (SP) 800-53 Rev. 5 (Withdrawn)). National Institute of Standards and Technology.
15. Force, J. T. (2022). Assessing security and privacy controls in information systems

and organizations. NIST Special Publication, 800, 53A.

16. Giraldo, J., Urbina, D., Cardenas, A., Valente, J., Faisal, M., Ruths, J., ... C Candell, R. (2018). A survey of physics-based attack detection in cyber-physical systems. ACM Computing Surveys (CSUR), 51(4), 1-36.
17. Goodfellow, I. J., Shlens, J., C Szegedy, C. (2014). Explaining and harnessing adversarial examples. arXiv preprint arXiv:1412.6572.
18. Greer, C., Wollman, D. A., Prochaska, D., Boynton, P. A., Mazer, J. A., Nguyen, C., ... C Bushby, S. T. (2014). NIST framework and roadmap for smart grid interoperability standards, release 3.0.
19. Prasanth Alluri. (2023, April). Privacy-Preserving Intrusion Detection in Pharmaceutical Information Systems Using Federated Learning. (JOCAA) : https://www.eudoxuspress.com/index.php/pub/article/view/4954/3712
20. Humayed, A., Lin, J., Li, F., C Luo, B. (2017). Cyber-physical systems security—A survey. IEEE Internet of Things Journal, 4(6), 1802-1831.
21. IEC. (2013). IEC 62443-3-3:2013 Industrial communication networks: Network and system security: System security requirements and security levels. International Electrotechnical Commission. https://webstore.iec.ch/en/publication/7033
22. IEC. (2018). IEC 62443-4-1: Security for industrial automation and control systems - Part 4-1: Secure product development lifecycle requirements.https://webstore.iec.ch/en/publication/33615
23. IEC. (2019). IEC 62443-4-2: Security for industrial automation and control systems - Part 4-2: Technical security requirements for IACS components.https://webstore.iec.ch/en/publication/34421
24. ISA. (2013). ANSI/ISA-62443-3-3-2013: Security for Industrial Automation and Control Systems: Part 3-3: System Security Requirements and Security Levels.International Society of Automation. https://www.isa.org/products/ansi-isa-62443- 3-3-2013-security-for-industrial-au
25. ISA. (2020). ANSI/ISA-62443-3-2-2020: Security risk assessment for system design. International Society of Automation. https://www.isa.org/products/ansi-isa-62443- 3-2-2020-security-for-industrial-a
26. ISO/IEC. (2022). ISO/IEC 27001:2022 Information security, cybersecurity and privacy protection: Information security management systems: Requirements. International Organization for Standardization. https://www.iso.org/obp/ui/en/#iso:std:iso-iec:27001:ed-3:v1:en
27. ISO/IEC. (2022). ISO/IEC 27002:2022 Information security, cybersecurity and privacy protection: Information security controls. International Organization for Standardization. https://www.iso.org/obp/ui/en/#iso:std:iso-iec:27002:ed-

3:v2:en

28. Johnson, A., Johnson, A., Dempsey, K., Ross, R., Gupta, S., C Bailey, D. (2011). Guide for security-focused configuration management of information systems. US Department of Commerce, National Institute of Standards and Technology.

29. Khedher, M. I.; Jmila, H.; Mounim A. El-Yacoubi. On the Formal Evaluation of the Robustness of Neural Networks and Its Pivotal Relevance for AI-Based Safety-Critical Domains. International Journal of Network Dynamics and Intelligence 2023, 2 (4), 100018. https://doi.org/10.53941/ijndi.2023.100018.

30. Kitsios, F., Chatzidimitriou, E., C Kamariotou, M. (2023). The ISO/IEC 27001 information security management standard: how to extract value from data in the IT sector. Sustainability, 15(7), 5828.

31. Knowles, W., Prince, D., Hutchison, D., Disso, J. F. P., C Jones, K. (2015). A survey of cyber security management in industrial control systems. International journal of critical infrastructure protection, 9, 52-80.

32. Kriaa, S., Pietre-Cambacedes, L., Bouissou, M., C Halgand, Y. (2015). A survey of approaches combining safety and security for industrial control systems. Reliability engineering C system safety, 139, 156-178.

33. Mathur, A. P., C Tippenhauer, N. O. (2016, April). SWaT: A water treatment testbed for research and training on ICS security. In 2016 international workshop on cyber- physical systems for smart water networks (CySWater) (pp. 31-36). IEEE.

34. MITRE. (2020). ATTCCK for ICS Matrix. The MITRE Corporation. https://attack.mitre.org/matrices/ics/

35. MITRE. (2021). MITRE ATLAS: Adversarial Threat Landscape for Artificial-Intelligence Systems. The MITRE Corporation. https://atlas.mitre.org/

36. MITRE. (2023). Best Practices for MITRE ATTCCK Mapping (including ICS considerations). U.S. Cybersecurity and Infrastructure Security Agency. https://www.cisa.gov/sites/default/files/2023-01/Best%20Practices%20for%20MITRE%20ATTCK%20Mapping.pdf

37. National Institute of Standards and Technology (US), C National Institute of Standards and Technology (US). (2024). NIST Cybersecurity Framework 2.0: Quick-start Guide for Creating and Using Organizational Profiles. US Department of Commerce, National Institute of Standards and Technology.

38. Prasanth Alluri. (2022). Behavior-Based Cyber Defense Architectures for Enhancing the Resilience of Defense and National Critical Infrastructure. Journal of Electrical Systems (JES) : https://journal.esrgroups.org/jes/article/view/9428

39. Nelson, A., Rekhi, S., Souppaya, M., C Scarfone, K. (2024). Incident response

recommendations and considerations for cybersecurity risk management: a CSF2.0 community profile (No. NIST Special Publication (SP) 800-61 Rev. 3 (Draft)). National Institute of Standards and Technology.

40. Papernot, N., McDaniel, P., Jha, S., Fredrikson, M., Celik, Z. B., C Swami, A. (2016, March). The limitations of deep learning in adversarial settings. In 2016 IEEE European symposium on security and privacy (EuroSCP) (pp. 372-387). IEEE.

41. Pillitteri, V. Y., C Brewer, T. L. (2014). Guidelines for smart grid cybersecurity.

42. Pillitteri, V. Y., Brewer, T. L., Feldman, L., C Witte, G. A. (2014). Release of NIST interagency report 7628 revision 1, guidelines for smart grid cybersecurity.

43. Ren, K., Zheng, T., Qin, Z., C Liu, X. (2020). Adversarial attacks and defenses in deep learning. Engineering, 6(3), 346-360.

44. Rose, S., Borchert, O., Mitchell, S., C Connelly, S. (2020). Zero trust architecture. NIST special publication, 800(207), 1-52.

45. Ross, R., Pillitteri, V., Dempsey, K., Riddle, M., C Guissanie, G. (2019). Protecting controlled unclassified information in nonfederal systems and organizations (No. NIST Special Publication (SP) 800-171 Rev. 2 (Draft)). National Institute of Standards and Technology.

46. Scarfone, K., Souppaya, M., Cody, A., C Orebaugh, A. (2008). Technical guide to information security testing and assessment. NIST Special Publication, 800(115), 2- 25.

47. Stouffer, K., Falco, J., C Scarfone, K. (2011). Guide to industrial control systems (ICS) security. NIST special publication, 800(82), 16-16.

48. Stouffer, K., Stouffer, K., Pease, M., Tang, C., Zimmerman, T., Pillitteri, V., ... C Thompson, M. (2023). Guide to operational technology (ot) security.

49. Syed, N. F., Shah, S. W., Shaghaghi, A., Anwar, A., Baig, Z., C Doss, R. (2022). Zero trust architecture (zta): A comprehensive survey. IEEE access, 10, 57143-57179.

50. Tuptuk, N., Hazell, P., Watson, J., C Hailes, S. (2021). A systematic review of the state of cyber-security in water systems. Water, 13(1), 81.

51. Prasanth Alluri. (2022). Integrating Artificial Intelligence for Climate-Resilient Energy Planning, Rural Infrastructure Development, and Water Conservation in Underdeveloped Regions. Computer Fraud and Security : https://computerfraudsecurity.com/index.php/journal/article/view/954

52. U.S. Department of Defense. (2022). DoD Zero Trust Strategy. Department of Defense Chief Information Officer.https://dodcio.defense.gov/Portals/0/Documents/Library/DoD-

ZTStrategy.pdf

53. U.S. Food and Drug Administration. (2023). Cybersecurity in Medical Devices: Quality System Considerations and Content of Premarket Submissions (Guidance for Industry and FDA Staff). U.S. Department of Health and Human Services.https://www.fda.gov/media/119933/download

54. Vinayakumar, R., Alazab, M., Soman, K. P., Poornachandran, P., Al-Nemrat, A., C Venkatraman, S. (2019). Deep learning approach for intelligent intrusion detection system. IEEE access, 7, 41525-41550.

55. Young, S. (2022). EXECUTIVE OFFICE OF THE PRESIDENT O F F I C E O F M A N A G E M E N T A N D B U D G E T MEMORANDUM FOR THE HEADS OF EXECUTIVEDEPARTMENTS AND AGENCIES FROM. https://www.whitehouse.gov/wp- content/uploads/2022/01/M-22-09.pdf