

A Multi-Layered Cybersecurity Management Framework for Robotics and Autonomous Systems in Industry 4.0 Manufacturing Environments

James U. Okoroama¹, Henry Olisakwe², Chiemela Onunka³

¹ Industrial Practitioner & Independent Researcher, MindCraft Synergy Consulting, United States

² Mechanical Engineering Department, Nnamdi Azikiwe University, Awka, Anambra State, Nigeria

³ Atlantic Constructors Inc., United States

Corresponding author: James U. Okoroama, james@mcssynergy.com

Abstract

Connecting industrial robots, autonomous mobile robots and safety controllers to enterprise and vendor networks has turned cyberattacks into potential sources of unsafe motion and injury, yet established cybersecurity frameworks are not written for robotic work cells. This conceptual paper develops the Multi-Layered Cybersecurity Framework for Robotics and Autonomous Systems. It derives a threat taxonomy linking spoofing, misconfiguration, supply chain compromise and insecure robot programming to physical consequences, maps five defensive layers to IEC 62443 foundational requirements and NIST Cybersecurity Framework categories, and proposes an ordinal, consequence-weighted risk index with a safety floor for prioritizing investment. The index is demonstrated on a constructed facility of six asset classes. Under the assumed parameters, the full layered program lowers the illustrative risk index by 73.7 percent, against 32.2 percent for a perimeter-only baseline, and brokered vendor access ranks as the most efficient first investment in 96.8 percent of 10,000 simulated parameter draws. A beta-factor extension shows how shared causes between layers erode the modelled benefit. These values describe model behavior under stated assumptions, not measured reductions in attack probability. The paper contributes an auditable structure for robotic cybersecurity programs and five propositions for empirical testing.

Keywords: Industrial cybersecurity; Robotics and autonomous systems; Industry 4.0; Zero trust architecture; Network segmentation; Incident response; Cyber-physical safety.

DOI: 10.21590/ijhit.06.4.319

1. Introduction

Manufacturing has moved through successive waves of automation, but the transition commonly described as Industry 4.0 differs from earlier waves in one decisive respect: it connects physical production equipment to digital networks that extend far beyond the factory wall (Lasi et al., 2014). Industrial robots, programmable logic controllers, autonomous mobile robots and machine vision systems now exchange data with manufacturing execution systems, enterprise resource planning platforms, cloud analytics services and vendor support portals. Lee et al. (2015) describe this arrangement as a cyber-physical systems architecture in which connection, conversion, cyber, cognition and configuration levels form a continuous stack. The productivity benefits are well established, including better utilization, condition-based maintenance and faster changeovers, but every new connection also creates a path through which an adversary can reach equipment that moves, cuts, lifts and welds (Xu et al., 2018).

The security consequences of this connectivity are qualitatively different from those in office information technology. In a business network, the dominant concerns are confidentiality and the integrity of records. In a robotic work cell, the dominant concerns are availability and the integrity of physical motion, because an altered trajectory, a disabled speed limit or a spoofed sensor reading can injure a technician standing within reach of a manipulator (Quarta et al., 2017). Humayed et al. (2017) note that cyber-physical systems inherit the vulnerabilities of both their cyber and physical components while adding new failure modes at the interface between them. The Stuxnet episode demonstrated more than a decade ago that a carefully engineered attack can manipulate industrial controllers while presenting operators with normal readings (Langner, 2011). Since then, the population of connected industrial devices has grown, and the attack surface has grown with it (Sisinni et al., 2018). Studies of IoT-driven critical infrastructure report the same expansion in transport and logistics settings, where connected devices now carry operational commands as well as telemetry (Jimoh et al., 2023a).

Robotic and autonomous systems sit at the sharp end of this problem. Experimental analysis of a widely deployed industrial robot controller showed that an attacker with network access could alter control parameters, tamper with calibration, change the robot's perceived state and suppress safety-related feedback to the operator (Quarta et al., 2017). Follow-up work found that many controllers were designed on the assumption that the network around them was trusted, and that

this assumption no longer holds once robots are connected for remote monitoring and maintenance (Pogliani et al., 2019). Autonomous mobile robots and drones add navigation and localization channels that can be spoofed; Kerns et al. (2014) showed that counterfeit satellite navigation signals could capture and redirect an unmanned aircraft.

Practitioners responsible for reliability, maintenance and operations in such facilities face a practical difficulty. A large body of guidance exists, including the NIST Cybersecurity Framework, the IEC 62443 series, NIST guidance on operational technology security and NIST guidance on zero trust architecture (National Institute of Standards and Technology [NIST], 2018; Rose et al., 2020; Stouffer et al., 2023). Each of these is valuable, yet none of them is written specifically for the robotic work cell, and none on its own tells a maintenance leader how to combine governance, access control, segmentation, detection and response into a single program that respects the safety constraints of moving machinery. Reviews of Industry 4.0 cybersecurity research have repeatedly found that the literature is fragmented across technical, organizational and economic perspectives, and that integrated management frameworks are scarce (Culot et al., 2019; Lezzi et al., 2018).

A second difficulty concerns how cybersecurity is framed inside industrial organizations. Gordon and Loeb (2002) showed that the economics of security investment depend on the vulnerability of the protected asset and the value at risk, which implies that investment should be targeted rather than spread evenly.

This paper addresses both difficulties. Its central research question is how manufacturing and logistics organizations can structure a layered cybersecurity management program for robotics and autonomous systems that integrates established standards, accounts for cyber-physical safety consequences and supports transparent investment prioritization. The paper makes four contributions. First, it consolidates a threat taxonomy for robotic and autonomous assets that links each threat class to its physical consequence. Second, it proposes the Multi-Layered Cybersecurity Framework for Robotics and Autonomous Systems (MLCF-RAS), which organizes defensive capability into five layers supported by cross-cutting enablers. Third, it introduces a simple quantitative model for inherent risk, residual risk and investment efficiency that maintenance and engineering managers can apply with data they already hold. Fourth, it

offers a phased implementation roadmap suitable for brownfield facilities in which legacy equipment cannot be replaced quickly.

Specifically, the paper addresses three research questions. RQ1 asks which threat classes affect robotic and autonomous assets in Industry 4.0 facilities and how each maps to physical consequence. RQ2 asks how established standards can be integrated into a layered management program that respects the safety constraints of moving machinery and the realities of legacy equipment. RQ3 asks how organizations can prioritize investment across assets and layers transparently while protecting safety-critical functions. These questions follow the design science logic of defining solution objectives before building an artifact (Peffer et al., 2007).

2. Background and Related Work

2.1 Convergence of Information Technology and Operational Technology

Historically, industrial control systems were engineered as closed environments. Programmable logic controllers, distributed control systems and robot controllers communicated over proprietary fieldbuses, and physical isolation was treated as the principal security control. Industry 4.0 dissolved much of this isolation by pushing sensor data upward to analytics platforms and pushing configuration, recipes and software updates downward from enterprise systems (Boyes et al., 2018). Sisinni et al. (2018) describe the Industrial Internet of Things as a set of technologies that connect machines, controllers and enterprise applications through standard internet protocols, which delivers interoperability but exposes devices that were built with minimal authentication and encryption. Data protection regulations differ markedly across jurisdictions, which complicates secure cloud implementation for operators of connected equipment (Mbonu et al., 2019a), and legal and ethical risk modeling has been proposed to bring these obligations into enterprise data protection governance (Mbonu et al., 2018). Data protection impact assessment models developed for multi-cloud infrastructure offer a structured way to evaluate exposure before new connections are made (Mbonu et al., 2022b).

Operational technology and information technology also differ in their priorities. Information technology security has traditionally emphasized confidentiality, integrity and availability in that order, whereas operational environments place availability and safety first because an unplanned stop or an uncontrolled motion carries immediate physical and economic consequences (Stouffer

et al., 2023). Patching cycles differ as well. Cherdantseva et al. (2016) found that risk assessment methods developed for business information systems frequently fail to capture these operational constraints, producing priorities that do not match plant reality.

Convergence also changes who touches the system. Ani et al. (2017) observed that the human and organizational dimensions of industrial cybersecurity, including skills, awareness and governance, are as important as technical controls, and that critical manufacturing infrastructure is particularly exposed because of the diversity of people and systems involved. Security architecture models for IT and OT convergence in regulated energy networks reach the same conclusion, arguing that design principles and governance alignment must be settled together (Adegbite et al., 2022).

A final feature of convergence is the growth of data flows that cross trust boundaries. Each hop is a potential point of interception or manipulation, and each new integration project tends to add flows faster than security teams can document them (Corallo et al., 2020). Cloud-hosted analytics and fleet services add a further layer of identity and configuration risk, since identity and access management must now span hybrid and multi-cloud estates (Mbonu et al., 2020a) and infrastructure-as-code templates can propagate an insecure setting to every deployment at once (Mbonu et al., 2020c).

2.2 Cyber-Physical Threats to Industrial Robotics and Autonomous Systems

The research literature on robot security has matured from proof-of-concept demonstrations to systematic analyses of attack surfaces. Clark et al. (2017) identified the principal categories of robotic cyber risk as compromise of the control software, manipulation of sensors and communication channels, and abuse of human-robot interaction interfaces. Quarta et al. (2017) defined an attacker model specific to industrial robots and demonstrated five classes of attack, ranging from control loop alteration to calibration tampering and user-perceived state manipulation, each of which could violate the basic requirements of accuracy, safety and integrity that a robot must satisfy. The authors also found that controllers were exposed on networks in ways that operators did not expect. Malware aimed at cloud-native and edge environments is increasingly classified with deep learning methods (Idika et al., 2021), and digital twin models of substations show how real-time monitoring can expose faults and anomalies in cyber-physical assets (Shittu et al., 2023). Threat actor analysis models help

organizations anticipate which adversaries are likely to target them and with which techniques (Dosunmu and Ogundele, 2023). Path-planning and slotting models for AMR-enabled warehouses show how tightly fleet movements are coupled to throughput (Akanbi and Sunday, 2024), which magnifies the operational consequence of a compromised fleet server.

Yaacoub et al. (2022) reviewed robotic cybersecurity across industrial, medical, military and service domains and concluded that robots combine the weaknesses of embedded systems, networked computers and autonomous decision systems. Their review highlighted weak authentication, unencrypted communication, outdated firmware, insecure default configurations and exposed debugging interfaces as recurrent problems. Pogliani et al. (2019) extended this analysis to the connected factory and argued that industrial robot safety standards had been written to address accidental failures rather than active adversaries, so that safety functions implemented in software could be undermined by an attacker who gained control of the same software environment.

Autonomous mobile robots and unmanned vehicles introduce further exposure through localization and navigation channels. Kerns et al. (2014) demonstrated that a spoofer could feed counterfeit navigation signals to an unmanned aircraft and steer it along an attacker-chosen trajectory without triggering alarms. Indoor mobile robots typically rely on lidar, wheel odometry, fiducial markers and wireless fleet management rather than satellite navigation, but the same principle applies: any input that the robot trusts without verification can be forged. Wu and Moon (2017) classified attacks on cyber-manufacturing systems by the domain in which they originate and the domain in which they produce effects, and showed that attacks crossing from the cyber domain into the physical domain are the most difficult to detect because their effects resemble ordinary equipment faults.

The broader industrial control system literature reinforces these findings. Pliatsios et al. (2020) surveyed supervisory control and data acquisition systems and documented the persistence of insecure legacy protocols that lack authentication. Mitchell and Chen (2014) reviewed intrusion detection for cyber-physical systems and noted that detection approaches built for enterprise networks perform poorly in environments where traffic is highly regular and where physical process models provide richer signals than packet contents. Recent reviews of machine learning for industrial control system anomaly detection support this emphasis while warning that learned

detectors can themselves be targeted by adversarial manipulation (Adebayo et al., 2022; Mitchell and Chen, 2014).

2.3 Existing Frameworks and Standards

Several established frameworks provide the building blocks for industrial cybersecurity programs. The NIST Cybersecurity Framework, Version 1.1, organizes cybersecurity activity into five functions, Identify, Protect, Detect, Respond and Recover, and is deliberately technology-neutral so that it can be tailored to sectors such as manufacturing (NIST, 2018). Version 2.0, released in February 2024, retains these functions and adds a sixth, Govern, which brings cybersecurity strategy, roles and supply chain risk management under explicit organizational oversight (NIST, 2024). Its limitation for robotic environments is that it does not prescribe how the functions should be implemented around safety-critical motion control, leaving organizations to interpret it for themselves. Comparative analyses of cybersecurity governance frameworks make the same observation, noting that frameworks provide structure but not sector-specific implementation detail (Jimoh et al., 2023b). Enterprise directory services remain a central dependency for these frameworks. Lessons from offline assessment of security-critical directory systems (Ladapo et al., 2018), from directory implementation in enterprise networks (Ladapo et al., 2019) and from analyses of directory attack steps and signatures (Ladapo et al., 2023a) show how identity infrastructure can become a single point of compromise. IT service management practice supplies the change and incident processes through which many controls are operated in practice (Ladapo et al., 2023b).

The IEC 62443 series addresses industrial automation and control systems specifically. IEC 62443-3-2 sets out a method for partitioning a system under consideration into security zones and conduits and for assigning target security levels based on risk, while IEC 62443-3-3 specifies system security requirements and capability levels (International Electrotechnical Commission [IEC], 2013, 2020). The zone and conduit model is especially relevant to robotic cells because it allows each cell, together with its safety controller and teach pendant, to be treated as a zone with tightly defined communication paths. NIST Special Publication 800-82, revised in 2023 as a guide to operational technology security, complements these standards with practical guidance on architecture, risk management and the application of security controls to operational systems (Stouffer et al., 2023).

Zero trust architecture represents a different line of development. NIST Special Publication 800-207 defines zero trust as a set of principles in which no implicit trust is granted to assets or users based solely on their network location, and in which access decisions are made per session using dynamic policy (Rose et al., 2020). Syed et al. (2022) surveyed zero trust implementations and found that while the model is maturing in enterprise environments, its adaptation to operational technology remains limited because many industrial devices cannot run modern authentication agents. Framework proposals for zero trust in operational technology have therefore emphasized phased implementation and resilience optimization rather than wholesale replacement of legacy assets (Ahmed et al., 2021), while identity-centric models place the identity of users and devices at the center of enterprise security governance (Ogbole et al., 2021). Incident handling guidance, notably NIST Special Publication 800-61, describes the lifecycle of preparation, detection and analysis, containment, eradication and recovery, and post-incident activity (Cichonski et al., 2012).

In parallel, defense in depth has long been recommended for industrial settings. Mosteiro-Sanchez et al. (2020) argued that defense in depth is the most suitable overarching strategy for Industry 4.0 because it tolerates the failure of any single control, and they combined it with lightweight end-to-end encryption for constrained industrial devices. Tuptuk and Hailes (2018) reviewed the security of smart manufacturing systems and identified the need for approaches that integrate security engineering into the lifecycle of manufacturing equipment rather than bolting it on afterwards.

2.4 Positioning Against Existing Frameworks

The contribution of the proposed framework is clearest when it is set against the frameworks on which it builds. Table 1 compares five established approaches with the framework developed in this paper across five dimensions that matter for robotic facilities: a threat taxonomy specific to robots and autonomous systems, explicit weighting of safety consequence, quantitative prioritization of investment, guidance for brownfield equipment that cannot run modern security agents, and treatment of workforce and supplier enablers. The established frameworks are strong on governance vocabulary (NIST, 2018), zone and conduit partitioning (IEC, 2020), per-session access decisions (Rose et al., 2020), operational technology architecture (Stouffer et al., 2023) and layered defense for the industrial internet of things (Mosteiro-Sanchez et al., 2020). None of

them, however, combines all five dimensions for robotic work cells, which is the gap the present framework is designed to fill.

Table 1. Comparison of established frameworks with the MLCF-RAS framework across five dimensions relevant to robotic facilities

Framework	Robot-specific threat taxonomy	Safety consequence weighting	Quantitative investment prioritization	Brownfield and legacy guidance	Workforce and supplier enablers
NIST Cybersecurity Framework 1.1 and 2.0	No	No	No	Partial	Partial
IEC 62443-3-2 and 3-3	No	Partial	Partial	Yes	Partial
NIST SP 800-207 zero trust	No	No	No	Partial	No
NIST SP 800-82 Rev. 3	Partial	Partial	No	Yes	Partial
Defence-in-depth for IIoT	No	No	No	Partial	No
MLCF-RAS (this paper)	Yes	Yes	Yes	Yes	Yes

2.5 Research Gap

Taken together, the literature offers strong technical analyses of robotic vulnerabilities and strong general-purpose frameworks for cybersecurity management, but the two have rarely been joined. Lezzi et al. (2018) reviewed cybersecurity research for Industry 4.0 and found that most studies addressed specific technical problems while few proposed integrated reference frameworks. Culot et al. (2019) reached a similar conclusion from an engineering management perspective and called for approaches that link security decisions to business impact. Elhabashy et al. (2019) developed a taxonomy of cyber-physical attacks on production systems from a quality control perspective and showed that conventional quality control was never designed to detect deliberate manipulation, which leaves organizations without a management view of how attacks propagate into physical outcomes. AI-enabled IT general controls (Mbonu et al., 2022a) and API governance with risk prioritization (Aliliele et al., 2023b) show the growing automation of control assurance, but these advances were developed mainly for financial and enterprise systems rather than robotic environments.

Three specific gaps motivate this paper. The first is the absence of a threat taxonomy that ties robotic threat classes to the physical, operational and financial consequences that matter to plant leaders. The second is the absence of guidance that integrates governance, zero trust, segmentation, detection and response into a coherent layered program for robotic cells, including legacy equipment that cannot support modern security agents. The third is the absence of a transparent prioritization method that maintenance and engineering managers can use to rank investments using data they already maintain in their asset registers and risk logs. The MLCF-RAS framework and its accompanying risk model are designed to address these gaps.

3. Methodology

This study follows a design science research approach, which is appropriate when the aim is to create and justify an artifact that solves an identified organizational problem (Hevner et al., 2004). Design science requires that the artifact be relevant to practice, grounded in existing knowledge and evaluated for its utility. Peffers et al. (2007) describe a six-activity process of problem identification, definition of solution objectives, design and development, demonstration, evaluation and communication, and this sequence structures the present work. The artifact is the MLCF-RAS framework together with its risk prioritization model and implementation roadmap. Maturity-model thinking from IT support operations (Anyanwu, 2019) informed the phased structure of the roadmap, and frameworks for digital transformation with embedded IT risk controls (Mbonu et al., 2020b) informed the treatment of governance as the foundation layer.

The knowledge base was assembled through an integrative review, an approach suited to topics whose evidence is spread across disciplines and where the aim is to synthesize a new conceptual model rather than to aggregate effects (Torraco, 2005; Snyder, 2019). Retained sources addressed threats to, or defenses for, industrial control, robotic or smart manufacturing systems, or offered a management framework adaptable to them, and were coded into threat mechanisms, consequence pathways, control families and implementation conditions; the taxonomy in Section 4 derives from the first two and the framework layers in Section 5 from the latter two.

The design was also informed by the first author's professional experience in reliability and maintenance engineering for automated industrial operations, which provided practical objectives for the artifact. In particular, solution objectives were defined as follows: the

framework should be usable by maintenance and engineering leaders as well as security specialists; it should accommodate legacy equipment; it should make cyber-physical safety consequences explicit; and it should support defensible prioritization of limited budgets. These objectives are consistent with risk-based assurance thinking, which ties control investment to the availability and integrity requirements of the assets concerned (Fadayomi et al., 2019). Demonstration takes the form of an illustrative application to a representative robotic facility in Section 6. Because the parameter values in that application are constructed for demonstration rather than measured at a specific site, the demonstration shows how the model operates rather than estimating effect sizes for any particular organization, a distinction consistent with the demonstration activity described by Peffers et al. (2007).

Because this is a conceptual article in the sense of Jaakkola (2020), its literature base was assembled to support theory synthesis rather than to estimate effects, and it is not a systematic review. Sources were selected purposively in two groups. The first comprises standards and peer-reviewed works on industrial control system security, robot security and zero trust, chosen because they define the concepts the framework integrates. The second comprises applied studies from a curated practitioner-research bibliography on enterprise and operational technology security, retained only where they addressed a specific framework element. Following the concept-centric approach recommended by Webster and Watson (2002), each source was coded to the framework element it informs, and Table 2 maps elements to sources so that the basis of each design choice can be traced. The review does not claim exhaustive coverage, and the selection was not accompanied by screening counts.

Table 2. Evidence map linking framework elements to the sources that inform them

Framework element	Principal conceptual sources	Applied and contextual sources	Use in the framework
Threat taxonomy	Quarta et al. (2017); Pogliani et al. (2019); Yaacoub et al. (2022)	Kerns et al. (2014); Wu and Moon (2017)	Defines threat classes and their physical consequences
Layer 1 governance and risk	NIST (2018); Kaplan and Garrick (1981); Cherdantseva et al. (2016)	Dosunmu and Ogundele (2019); Fadayomi et al. (2019)	Structures governance functions and the risk index
Layer 2 zero trust access	Rose et al. (2020); Saltzer and Schroeder (1975); Syed et al. (2022)	Ahmed et al. (2021); Oshoba et al. (2019)	Defines brokered access and least privilege for legacy devices
Layer 3 segmentation	IEC (2020); Stouffer et al. (2023)	Mosteiro-Sanchez et al. (2020)	Defines zones, conduits and safety network separation

Framework element	Principal conceptual sources	Applied and contextual sources	Use in the framework
Layer 4 detection	Mitchell and Chen (2014); Pliatsios et al. (2020)	Dosunmu and Ogundele (2020); Ladapo et al. (2022a)	Combines network and process-aware monitoring
Layer 5 response and recovery	Cichonski et al. (2012); Alcaraz and Zeadally (2015)	Ojukwu et al. (2023); Dosunmu and Ogundele (2021)	Adapts the incident lifecycle to safe states and golden images

4. Threat Landscape for Robotic and Autonomous Work Cells

The threat taxonomy developed in this section organizes the principal threat classes affecting robotic and autonomous assets according to their mechanism, typical entry point and dominant consequence. Four classes recur across the reviewed literature: spoofing and sensor manipulation, misconfiguration, supply chain compromise, and insecure programming or weakened safety logic. These classes are not mutually exclusive, since a single campaign may exploit a misconfigured remote access service to deliver a tampered program, but separating them clarifies which defensive layer is best placed to interrupt each one (Wu and Moon, 2017). Table 3 summarizes the taxonomy, and the subsections that follow discuss each class.

Table 3. Threat taxonomy for robotics and autonomous systems in Industry 4.0 facilities

Threat class	Mechanism	Typical entry point	Dominant consequence
Spoofing and sensor manipulation	Impersonation of a device, user or signal; forged localization or vision input	Wireless fleet network, sensor buses, navigation channels	Unsafe motion, collision, misrouted vehicles
Misconfiguration	Default credentials, exposed services, flat networks, permissive cloud settings	Engineering workstations, remote access services, cloud tenants	Unauthorized access, lateral movement, loss of control
Supply chain compromise	Tampered firmware, libraries, robot applications or integrator tools	Vendor updates, third-party packages, integrator laptops	Persistent hidden access, widespread disruption
Insecure programming and weakened safety logic	Unsafe robot code patterns, altered speed and force limits, bypassed interlocks	Teach pendant, offline programming tools, controller file systems	Injury risk, product defects, equipment damage

4.1 Spoofing and Sensor Manipulation

Spoofing is an attack in which an adversary impersonates a legitimate device, user or signal in order to steal data, deliver malicious instructions or circumvent access restrictions. In robotic environments, spoofing extends from network identity to the physical inputs on which

autonomous behavior depends. Kerns et al. (2014) showed that a navigation spoofer could gradually pull an unmanned aircraft off course while the vehicle's own estimator accepted the forged signal as genuine.

Network-level spoofing is equally relevant. Many industrial protocols lack message authentication, so an attacker who reaches the control network can impersonate a supervisory controller and issue commands that the robot will execute (Pliatsios et al., 2020). Quarta et al. (2017) demonstrated that manipulating the state information displayed to the operator could conceal an attack, so that a technician believes the robot is in a safe mode when it is not. This form of deception is particularly dangerous during maintenance, when people work inside the cell under the assumption that interlocks and reduced-speed modes are active. The defensive implication is that both commands and state reports must be verified, and that physical safety functions must remain independent of the networked control path. Countermeasures such as attribute-based encryption for access control in connected supply chain systems illustrate how message-level verification can be added without redesigning every endpoint (Shittu and Nabil, 2023).

4.2 Misconfiguration

Misconfiguration remains one of the most common root causes of industrial security incidents because complex environments contain thousands of settings, and a single permissive rule can undo the protection provided by many correct ones. Typical examples include default or shared credentials on robot controllers and human-machine interfaces, unused services left enabled, flat networks in which robot cells can communicate freely with office systems, and remote access tools installed for commissioning and never removed (Stouffer et al., 2023). As robotic facilities adopt cloud-hosted fleet management and analytics, misconfiguration also extends to cloud tenants, identity providers and application programming interfaces. Blockchain-enabled audit trails for cloud configuration management have been proposed to make configuration changes tamper-evident (Oshoba et al., 2020), and cloud migration practice stresses that security baselines must be defined before workloads move (Ladapo et al., 2022b).

Misconfiguration in robotic systems has a distinctive character because it frequently originates in commissioning and maintenance practices rather than in information technology administration. Pogliani et al. (2019) observed that industrial robots are often deployed with network services

exposed beyond what operation requires, partly because security requirements are not included in commissioning acceptance tests. Kolias et al. (2017) documented how the Mirai botnet exploited default credentials on internet-connected devices at enormous scale, illustrating how quickly a known weakness can be harvested once it is widespread. For robotic facilities, the defensive implication is that configuration baselines must become part of maintenance standards, verified after every intervention in the same way that mechanical torque settings or safety circuit tests are verified. Continuous cloud misconfiguration monitoring frameworks show how such verification can be automated for the cloud components of robotic fleets (Aliliele et al., 2023a), and the same discipline extends to the access governance of large cloud tenants that host fleet management services (Mbonu et al., 2022c).

4.3 Supply Chain Compromise

Robotic systems depend on long supply chains that include robot manufacturers, controller and drive suppliers, vision and sensor vendors, systems integrators and providers of software libraries and fleet management platforms. Each link can introduce compromised components. Ohm et al. (2020) reviewed attacks on open-source software supply chains and showed that malicious packages are inserted through techniques such as typosquatting, account takeover and injection into build processes, after which they propagate to every downstream user. Robot applications increasingly incorporate such packages, particularly in vision, path planning and fleet coordination software.

Supply chain risk also arises from privileged vendor access. Original equipment manufacturers commonly require remote connectivity to support diagnostics, warranty service and software updates, and integrators often carry engineering laptops between customer sites. A compromise of the vendor environment can therefore become a compromise of every customer facility that trusts it (Ani et al., 2017). Abomhara and Køien (2015) noted that the trust relationships embedded in connected device ecosystems are often implicit and poorly documented, which makes them attractive to attackers. The defensive implications are that vendor access must be brokered, time-limited and monitored; that software and firmware should be verified before installation; and that procurement contracts should specify security obligations, including vulnerability disclosure and update support across the expected service life of the equipment.

Insider threat models are also relevant here, because vendor and integrator personnel hold trusted access that can be abused deliberately or carelessly (Akeju et al., 2018).

4.4 Insecure Programming and Weakened Safety Logic

Industrial robots are programmed through vendor-specific languages that, in addition to motion instructions, often provide access to files, network sockets and system resources. These capabilities are useful for integration but allow unsafe patterns, such as programs that accept motion targets from the network without validation or that can be modified at run time. Quarta et al. (2017) showed that altering configuration and program parameters could change a robot's accuracy and violate safety constraints while remaining difficult for operators to notice. Artificial intelligence techniques for secure software testing and automated regression control (Mbonu et al., 2021), and broader applications of artificial intelligence in software testing (Borah et al., 2022), can be adapted to vet robot programs before they are commissioned.

The most serious consequence of this threat class is the weakening of safety logic. Collaborative and fenceless robotic applications rely on speed and separation monitoring, power and force limiting and safety-rated monitored stops to protect people who share the workspace (Villani et al., 2018). Clark et al. (2017) emphasized that robots are physical actors and that attacks on their control logic can translate directly into harm. The defensive implications are that safety-rated functions should be implemented in independent, certified safety controllers wherever possible; that changes to robot programs and safety parameters should be subject to authenticated change management; and that program integrity should be verified routinely against approved baselines. Privacy-centric and security-by-design engineering approaches developed for digital and cloud systems offer a template for building these checks into the robot software lifecycle rather than adding them after commissioning (Okoruwa et al., 2020).

5. The Multi-Layered Cybersecurity Framework for Robotics and Autonomous Systems

The MLCF-RAS framework organizes defensive capability into five layers that surround the protected robotic and autonomous assets, supported by a set of cross-cutting enablers. The layering follows the defense in depth principle that each control should be assumed fallible and that an attacker should have to defeat several independent barriers before reaching a safety-

critical function (Mosteiro-Sanchez et al., 2020). Figure 1 presents the structure. The cross-cutting enablers, shown as a vertical column, influence every layer simultaneously.

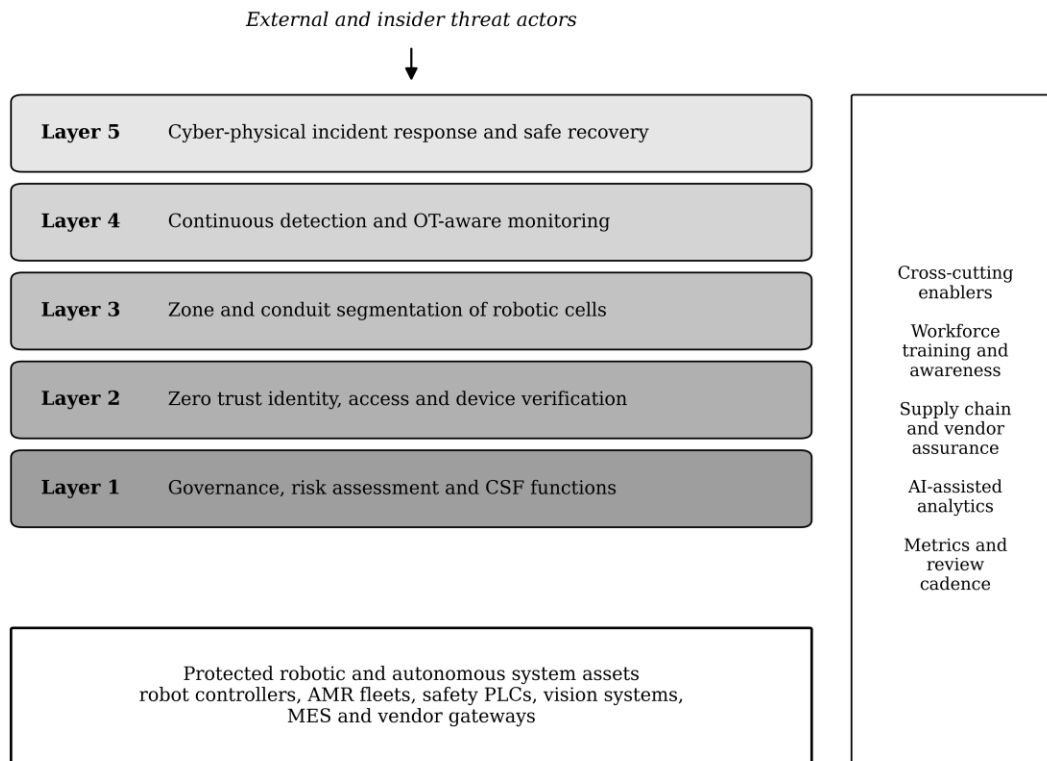


Figure 1. Structure of the MLCF-RAS framework, showing five defensive layers around robotic and autonomous system assets and four cross-cutting enablers.

Figure 1 should be read from the asset outward. The layers are deliberately complementary rather than alternative, a principle consistent with the NIST Cybersecurity Framework's view that its functions operate concurrently and continuously (NIST, 2018).

5.1 Layer 1: Governance, Risk Assessment and Cybersecurity Framework Functions

The foundation of the framework is governance, which establishes accountability, scope and risk appetite for robotic cybersecurity. The NIST Cybersecurity Framework provides the organizing structure because its Identify, Protect, Detect, Respond and Recover functions offer a vocabulary shared by engineers, executives and auditors and because it can be tailored through profiles to the manufacturing sector (NIST, 2018). Under the Identify function, the first practical task is a complete inventory of robotic and autonomous assets, including controllers, teach pendants, safety controllers, mobile robots, vision systems, fleet servers, engineering workstations and

every remote access path. Security audit and enterprise risk assessment frameworks provide the structure for closing this gap systematically (Dosunmu and Ogundele, 2019).

Risk assessment is the second governance task. Cherdantseva et al. (2016) found that many risk assessment methods for industrial control systems rely heavily on expert judgment and seldom incorporate physical consequence explicitly. The MLCF-RAS framework therefore adopts the consequence-weighted model developed in Section 6, which scores safety, operational and financial impact separately. This keeps human safety visible in prioritization rather than letting it be absorbed into a single loss estimate. The classical definition of risk as a set of scenarios, each with a likelihood and a consequence, provides the conceptual basis (Kaplan and Garrick, 1981).

Governance also requires clear ownership across functions that rarely share reporting lines. In many facilities, information security belongs to a corporate technology function, while robots belong to engineering and maintenance. The framework assigns a named owner for each robotic zone, typically the reliability or controls engineering lead, with security specialists acting as advisers and auditors. Reason (1997) argued that organizational accidents emerge from latent conditions created by management decisions; cybersecurity governance in robotic facilities is a way of identifying and removing such latent conditions before an attacker exploits them.

5.2 Layer 2: Zero Trust Identity, Access and Device Verification

The second layer applies zero trust principles to the people, devices and applications that interact with robotic systems. Zero trust rejects the idea that a connection is trustworthy because it originates inside the plant network and instead requires that each access request be authenticated, authorized and evaluated against policy at the time it is made (Rose et al., 2020). This principle is directly relevant to the threat classes in Section 4, since spoofing, misconfiguration and supply chain attacks all exploit implicit trust. The older principles of least privilege and complete mediation articulated by Saltzer and Schroeder (1975) remain the conceptual core: every subject should operate with the minimum privileges required, and every access should be checked. Secure identity and access management models for distributed and federated systems (Oshoba et al., 2019) and reviews of access control policies grounded in privacy requirements (Ladapo et al., 2023c) provide design patterns for brokered access to robot controllers. Secure cloud architectures for enterprise messaging show how such brokers can scale without weakening authentication (Ahmed and Odejobi, 2018).

Implementing zero trust around robots requires adaptation, because many controllers cannot run modern authentication agents. The framework therefore places policy enforcement points in front of legacy devices rather than on them. Engineering access to robot controllers is routed through a hardened jump host or secure remote access broker that enforces multi-factor authentication, records sessions and grants time-limited permissions tied to an approved work order. Syed et al. (2022) identified this proxy-based pattern as the most practical route to zero trust where endpoints are constrained. Device identity is established through certificates for newer equipment and through network access control and asset fingerprinting for older equipment.

Role design is equally important. Operators, maintenance technicians, robot programmers, integrators and vendor engineers require different privileges, and these privileges should map to tasks rather than to people's seniority. Changes to safety-related settings should require dual authorization by a qualified engineer and, where feasible, confirmation at the physical controller. Bulgurcu et al. (2010) showed that employees' compliance with security policies depends on their beliefs about the costs and benefits of compliance, so access workflows must be designed to be fast enough that technicians do not seek workarounds during breakdowns.

5.3 Layer 3: Zone and Conduit Segmentation of Robotic Cells

The third layer divides the network into zones of similar security requirement and permits communication between zones only through defined conduits. This approach, formalized in IEC 62443-3-2, limits the spread of an intrusion and makes abnormal communication easier to detect (IEC, 2020). In the MLCF-RAS framework each robotic cell, together with its controller, safety controller, teach pendant and local vision system, forms its own zone. Mobile robot fleets form a separate zone with their fleet management server.

Conduits between zones should permit only the protocols and directions that operation requires. A robotic cell typically needs to receive production orders and send status and quality data, but it rarely needs to initiate connections to the enterprise network or the internet. Stouffer et al. (2023) recommend that operational networks be separated from business networks by firewalls configured to deny traffic by default, with exceptions documented and reviewed. Within the cell, the safety network should be logically or physically separate from the general control network so that a compromise of non-safety functions cannot reach safety-rated components.

Segmentation delivers value even before it is complete. Mosteiro-Sanchez et al. (2020) noted that segmentation combined with encryption of data in transit provides protection that persists even when intermediate devices are compromised. For brownfield sites, segmentation can often be achieved with industrial firewalls inserted at cell boundaries without modifying the robots themselves, which makes it one of the most practical controls available for legacy equipment. Experience with operational technology networks in electric utilities under NERC CIP requirements indicates that electronic security perimeters around critical assets deliver substantial risk reduction even when the assets themselves remain unchanged (Stouffer et al., 2023).

5.4 Layer 4: Continuous Detection and Operational Technology Aware Monitoring

The fourth layer assumes that preventive controls will sometimes fail and focuses on detecting intrusions and anomalous behavior quickly. Detection in robotic environments benefits from the regularity of industrial traffic. Robots, controllers and supervisory systems communicate in predictable patterns, so deviations such as a new device, a new protocol, an unexpected program download or a configuration change outside a maintenance window stand out clearly against the baseline (Mitchell and Chen, 2014). Passive network monitoring tools that understand industrial protocols can build this baseline without placing load on fragile devices. Reviews of SIEM, network detection and response, and anomaly detection architectures for ICS and SCADA environments describe how these sources can be combined (Pliatsios et al., 2020), and intrusion detection and prevention models show how detection outputs translate into organizational defense effectiveness (Dosunmu and Ogundele, 2020). Threat intelligence integration further sharpens detection by telling analysts which techniques are currently in use against comparable facilities (Dosunmu and Ogundele, 2022). Human-in-the-loop machine learning keeps analysts involved in labeling and validating alerts, which improves detection where attacks are rare (Ladapo et al., 2022a). Reliability modeling of maintenance-free wireless infrastructure is relevant where robot telemetry travels over sensor networks whose own availability must be monitored (Asiedu and Quainoo, 2023). Breach and attack simulation offers a way to verify continuously that detection rules actually fire against realistic attack paths (Dosunmu and Ogundele, 2024a). Energy-harvesting sensors, such as the piezoelectric systems studied for self-powered roadside equipment (Ejimah and Aregbe, 2022), can extend monitoring to locations

where cabling is impractical, but each such device is itself an endpoint that must be inventoried and secured.

Process-aware monitoring adds a second detection channel. Because robotic attacks often manifest as physical anomalies, signals such as cycle time, joint torque, motor current, path deviation and vision rejection rates can reveal tampering that network monitoring misses. Wu and Moon (2017) argued that cross-domain attacks are hard to distinguish from ordinary faults, which implies that maintenance and reliability data should be shared with security analysts. In practice this means that condition monitoring systems used for predictive maintenance can also serve as security sensors, provided that their alerts are correlated with network and access logs.

Detection also depends on people. Parsons et al. (2014) showed that employees' security awareness can be measured and improved, and that knowledge and attitude both influence behavior.

5.5 Layer 5: Cyber-Physical Incident Response and Safe Recovery

The fifth layer prepares the organization to respond when an incident occurs. Because it is unlikely that any organization can defend against every attack, preparation for response is as important as prevention. NIST guidance describes the incident lifecycle as preparation, detection and analysis, containment, eradication and recovery, and post-incident activity (Cichonski et al., 2012). In robotic environments this lifecycle must be adapted so that the first priority of containment is human safety. When a robot may be compromised, the response plan should specify how to bring it to a safe state using hardwired emergency stops and safety controllers that do not depend on the potentially compromised control network. Adaptive protection schemes for digital substations show how IoT-enabled cyber-physical resilience can be engineered into critical equipment rather than added afterwards (Shittu et al., 2024).

Recovery must also respect the validation requirements of robotic cells. Restoring a controller from backup is insufficient if the backup itself may contain tampered programs, so golden images of programs, configurations and safety parameters should be maintained offline, version-controlled and hashed. Before a cell returns to production, its safety functions should be functionally tested in the same way they are tested after mechanical maintenance. Alcaraz and Zeadally (2015) emphasized that resilience in critical infrastructure depends on the ability to

restore service rapidly and safely, which requires that recovery procedures be practiced rather than merely documented. Ransomware targeting operational technology has made this requirement urgent, and physics-informed detection combined with digital twins and resilient recovery procedures has been proposed as a defense suited to cyber-physical plants (Ojukwu et al., 2023). Blockchain-assisted data exchange has likewise been proposed to protect the integrity of records that operators rely on during recovery (Shittu et al., 2022).

Post-incident reviews should feed lessons back into the governance layer so that the framework improves continuously, mirroring the learning cycle that reliability organizations already apply to equipment failures through root cause analysis (Reason, 1997). Incident response and digital forensics strategies for rapid containment (Dosunmu and Ogundele, 2021) and automated log analytics linked to response playbooks (Mbonu et al., 2019b) give the post-incident review the evidence it needs.

5.6 Cross-Cutting Enablers

Four enablers cut across all five layers. The first is workforce training and awareness. A well-trained workforce is a critical component of any cybersecurity program because people configure, maintain and operate the systems that controls are meant to protect. Bulgurcu et al. (2010) found that awareness shapes employees' beliefs about the consequences of compliance and non-compliance, and Ani et al. (2017) identified human capability as a persistent weakness in industrial infrastructure. Studies of password practices and authentication behaviour among professional staff show that everyday credential habits remain a persistent weakness even in well-governed organizations (Onche and Ogbonna, 2022). Training for robotic environments should be role-specific: operators need to recognize and report anomalies, technicians need to follow secure maintenance procedures, and engineers need to understand secure programming and configuration management. Awareness research shows that confidentiality practices depend on regular reinforcement (Onche et al., 2020), and personal device use introduces risks that also apply to technicians who carry their own phones and tablets onto the floor (Onche et al., 2023). Blockchain-based cross-border data exchange illustrates how trust between partners can be engineered when operational data must move between organizations (Kumuyi et al., 2023).

The second enabler is supply chain and vendor assurance. Protecting the supply chain prevents the introduction of compromised components and limits the exposure created by privileged third parties (Ohm et al., 2020).

The third enabler is analytics, including artificial intelligence and machine learning. These techniques can strengthen detection by learning normal patterns of network traffic and process behavior, but they can also be used by adversaries, so organizations should adopt them with attention to their limitations and to the explainability of their alerts (Mitchell and Chen, 2014). The fourth enabler is measurement. Gordon and Loeb (2002) showed that security investment should be calibrated to the value of the assets protected; regular measurement allows leaders to see whether investment is reducing risk where it matters most.

5.7 Mapping of Layers to Threats and Standard Provisions

To make the framework auditable against existing obligations, Table 4 maps each layer to the threat classes it primarily addresses, to the foundational requirements of IEC 62443-3-3 and to categories of the NIST Cybersecurity Framework. The seven foundational requirements of IEC 62443-3-3 are identification and authentication control, use control, system integrity, data confidentiality, restricted data flow, timely response to events and resource availability (IEC, 2013), and the mapping shows that the five layers jointly cover all seven. The mapping also shows that governance and response, which standards treat as program functions, carry robotic-specific content, such as configuration baselines for robot programs and safe-state procedures for compromised cells (NIST, 2018). Categories are given in the notation of version 1.1, which remains widely used in operational technology programs; in version 2.0, supply chain risk management moves to the Govern function, which corresponds directly to Layer 1 of the framework (NIST, 2024).

Table 4. Mapping of MLCF-RAS layers to threat classes, IEC 62443-3-3 foundational requirements and NIST CSF 1.1 categories

Layer	Threat classes primarily addressed	IEC 62443-3-3 foundational requirements	NIST CSF 1.1 categories
1. Governance and risk	Misconfiguration; supply chain compromise; insecure programming	FR3 System integrity (configuration baselines)	ID.AM Asset Management; ID.RA Risk Assessment; ID.SC Supply Chain Risk Management; PR.IP Information Protection Processes

Layer	Threat classes primarily addressed	IEC 62443-3-3 foundational requirements	NIST CSF 1.1 categories
2. Zero trust access	Spoofing; misconfiguration; supply chain (vendor access)	FR1 Identification and authentication control; FR2 Use control	PR.AC Identity Management and Access Control
3. Segmentation	All classes (containment)	FR5 Restricted data flow; FR4 Data confidentiality	PR.AC-5 network integrity; PR.PT Protective Technology
4. Detection	Spoofing; insecure programming; supply chain	FR6 Timely response to events; FR3 System integrity	DE.AE Anomalies and Events; DE.CM Security Continuous Monitoring
5. Response and recovery	All classes (consequence limitation)	FR7 Resource availability; FR6 Timely response to events	RS.RP Response Planning; RS.MI Mitigation; RC.RP Recovery Planning

6. Quantitative Risk Prioritization Model

Layered frameworks can become unwieldy if every control is pursued with equal urgency. Managers need a way to decide which assets to protect first and which layers deliver the greatest reduction in risk for the money spent. Table 5 first shows how the five layers address the four threat classes identified in Section 4, distinguishing primary from supporting contributions. The pattern is consistent with the defense in depth logic: no threat class depends on a single layer, and every layer contributes to at least three threat classes (Mosteiro-Sanchez et al., 2020). The coverage matrix is qualitative, however, and the model developed in this section translates it into numbers that can support budgeting.

Table 5. Coverage of threat classes by MLCF-RAS layers (P = primary contribution, S = supporting contribution)

Threat class	Layer 1 Governance	Layer 2 Zero trust	Layer 3 Segmentation	Layer 4 Detection	Layer 5 Response
Spoofing and sensor manipulation	S	P	S	P	S
Misconfiguration	P	P	S	S	S
Supply chain compromise	P	P	S	P	S
Insecure programming and weakened safety logic	P	P	S	P	P

6.1 Model Formulation

The model draws on the classical definition of risk as the combination of the likelihood of a scenario and the magnitude of its consequence (Kaplan and Garrick, 1981). For each robotic

asset or asset class i , the inherent risk before the layered controls are applied is defined in Equation (1) as the product of a likelihood score and a composite consequence score. Cyber risk quantification models developed for enterprise investment decisions pursue the same goal of ranking security spending by risk reduction (Dosunmu and Ogundele, 2024b), although they rarely weight physical safety consequences as this model does.

$$R_i^0 = L_i \cdot C_i \quad (1)$$

In Equation (1), L_i is the likelihood of a successful attack on asset i , scored from 1 (rare) to 5 (almost certain) on the basis of exposure, known vulnerabilities and attacker interest, and C_i is the composite consequence. Because robotic incidents carry several distinct kinds of consequence, C_i is computed in Equation (2) as a weighted sum of three consequence dimensions.

$$C_i = \sum_{k=1}^3 w_k c_{ik} \quad (2)$$

Here c_{i1} , c_{i2} and c_{i3} are the safety, operational and financial consequence scores of asset i , each scored from 1 to 5, and w_k are the weights assigned to each dimension. The weights must satisfy the conditions in Equation (3).

$$\sum_{k=1}^3 w_k = 1, w_k \geq 0 \quad (3)$$

Separating the dimensions follows the observation by Cherdantseva et al. (2016) that industrial risk methods often collapse physical consequence into financial loss, hiding the safety stakes. In this application the safety weight is set to 0.5, the operational weight to 0.3 and the financial weight to 0.2, reflecting a policy that human safety dominates. Organizations can adjust these weights to reflect their own risk appetite, but the weights should be set by leadership before scoring begins so that they are not tuned to justify a preferred conclusion.

Each layer of the framework reduces risk by an effectiveness fraction e_{ij} between 0 and 1, which represents the proportion of remaining risk that layer j removes for asset i . Because the layers act in sequence on whatever risk the previous layers leave behind, residual risk is obtained multiplicatively, as shown in Equation (4).

$$R_i^{res} = R_i^0 (1 - e_{i1})(1 - e_{i2}) \cdots (1 - e_{i5}) \quad (4)$$

The multiplicative form captures the defense in depth property that each layer acts on residual rather than original risk, so that the combined effect of several moderately effective layers can

exceed that of a single strong one (Mosteiro-Sanchez et al., 2020). The proportional risk reduction for each asset follows from Equation (5).

$$\Delta R_i = 1 - \frac{R_i^{res}}{R_i^0} \quad (5)$$

Finally, to support budgeting, the model computes an investment efficiency index in Equation (6), where K_i is the cost of implementing the five layers for asset i . The index expresses the absolute risk reduction obtained per unit of cost and therefore indicates where the next unit of budget yields the most protection, consistent with the investment logic described by Gordon and Loeb (2002).

$$\pi_i = \frac{R_i^0 - R_i^{res}}{K_i} \quad (6)$$

6.2 Illustrative Application

The model is demonstrated on an illustrative robotic fulfillment and assembly facility containing six asset classes: articulated robot controllers with networked teach pendants; safety programmable logic controllers and interlock circuits; an autonomous mobile robot fleet management server; machine vision and guidance cameras; the gateway linking the manufacturing execution system to enterprise resource planning; and the vendor remote access gateway. The parameter values were constructed to represent a typical brownfield facility and are intended to show how the model behaves, consistent with the demonstration activity in design science research (Peffers et al., 2007). They are not measurements from a particular site. Table 6 reports the likelihood and consequence scores together with the resulting inherent risk.

Table 6. Illustrative likelihood, consequence and inherent risk scores for six robotic asset classes

Asset class	L	Safety c1	Operational c2	Financial c3	C	Inherent risk R0
Robot controller and teach pendant	4	5	4	4	4.50	18.00
Safety PLC and interlock circuit	2	5	5	3	4.60	9.20
AMR fleet management server	4	4	5	3	4.10	16.40
Machine vision and guidance camera	3	4	3	3	3.50	10.50
MES and ERP integration gateway	3	2	4	5	3.20	9.60
Vendor remote access gateway	5	4	4	4	4.00	20.00

As Table 6 shows, the vendor remote access gateway carries the highest inherent risk because it combines high exposure with significant consequence across all three dimensions. The robot controller and the mobile robot fleet server follow closely. The safety controller has the highest composite consequence of all assets, but its likelihood score is lower because, in the illustrative facility, it sits on a separate safety network with limited external connectivity. This illustrates why consequence and likelihood must be assessed separately: an asset can be critical without being the most exposed (Kaplan and Garrick, 1981).

The layer effectiveness values in Table 7 were set to reflect how each layer acts on each asset class. Governance contributes a modest, uniform reduction because its effect is indirect. Zero trust is most effective for assets accessed by people and external parties, such as the vendor gateway, while segmentation is most effective for the safety controller, which benefits most from strict isolation. Detection is most effective for assets whose behavior generates rich process signals.

Table 7. Assumed layer effectiveness values $e(ij)$ for the illustrative facility

Asset class	Layer 1	Layer 2	Layer 3	Layer 4	Layer 5
Robot controller and teach pendant	0.10	0.30	0.35	0.20	0.15
Safety PLC and interlock circuit	0.10	0.25	0.45	0.15	0.15
AMR fleet management server	0.10	0.35	0.30	0.25	0.15
Machine vision and guidance camera	0.10	0.20	0.30	0.25	0.10
MES and ERP integration gateway	0.10	0.35	0.25	0.25	0.15
Vendor remote access gateway	0.10	0.50	0.30	0.25	0.15

6.3 Results of the Illustrative Application

Applying Equations (4) to (6) to the values in Tables 6 and 7 yields the residual risk, proportional reduction and investment efficiency reported in Table 8. Illustrative implementation costs, expressed in thousands of United States dollars, are included to compute the efficiency index.

Table 8. Residual risk, risk reduction and investment efficiency for the illustrative facility

Asset class	Inherent risk R_0	Residual risk	Reduction (%)	Cost K (USD thousand)	Efficiency index	Rank by efficiency
Vendor remote access gateway	20.00	4.02	79.9	30	0.533	1

Asset class	Inherent risk R0	Residual risk	Reduction (%)	Cost K (USD thousand)	Efficiency index	Rank by efficiency
Machine vision and guidance camera	10.50	3.57	66.0	25	0.277	2
AMR fleet management server	16.40	4.28	73.9	55	0.220	3
Robot controller and teach pendant	18.00	5.01	72.2	60	0.216	4
Safety PLC and interlock circuit	9.20	2.47	73.2	35	0.192	5
MES and ERP integration gateway	9.60	2.69	72.0	45	0.154	6

Across the six asset classes, total risk falls from 83.7 to 22.0 points, an aggregate reduction of 73.7 percent. Individual reductions range from 66.0 percent for the vision camera to 79.9 percent for the vendor gateway, as shown in Table 8. No single layer achieves these results; for every asset, the largest single-layer effectiveness is 0.50 or less, and the combined reduction arises from stacking several moderate controls. This is the quantitative expression of the defense in depth principle that motivates the framework (Mosteiro-Sanchez et al., 2020).

The efficiency ranking in Table 8 carries a practical message. The vendor remote access gateway delivers more than twice the risk reduction per unit of cost of any other asset, because a relatively inexpensive brokered access solution addresses a highly exposed entry point. This finding is consistent with the supply chain analysis in Section 4.3 and with the broader observation that third-party access is a frequent initial vector in industrial incidents (Ani et al., 2017). The vision camera ranks second because its controls are inexpensive, even though its inherent risk is moderate.

The safety controller ranks fifth by efficiency, which illustrates an important limitation of purely economic prioritization. Its residual risk is the lowest in the facility, but its safety consequence score is the maximum, and a failure would expose people to harm. For this reason, the MLCF-RAS framework applies a consequence floor: any asset with a safety consequence score of 5 is protected to a defined minimum standard regardless of its efficiency rank. Efficiency then governs the sequencing of the remaining investments. This rule reflects the principle, long established in safety management, that organizations should not trade the protection of people against marginal financial returns (Reason, 1997).

6.4 Monte Carlo Robustness Analysis

Because the illustrative parameters in Tables 6 and 7 were set by judgment, it is important to show that the main conclusions do not depend on their exact values. A Monte Carlo analysis with 10,000 draws was therefore performed. In each draw, every likelihood score was shifted by minus one, zero or plus one within the range of 1 to 5; the consequence weights were drawn from a Dirichlet distribution centered on the policy weights of 0.5, 0.3 and 0.2; every layer effectiveness value was multiplied by a factor drawn uniformly between 0.7 and 1.3; and every implementation cost was multiplied by a factor between 0.8 and 1.2. Probabilistic sensitivity analysis of this kind is recommended where risk models depend on expert judgment (Cherdantseva et al., 2016). Table 9 reports the results.

Table 9. Monte Carlo robustness of residual risk reduction and investment ranking (10,000 draws)

Asset class	Mean risk reduction (%)	5th percentile (%)	95th percentile (%)	Probability ranked first by efficiency (%)	Probability in top three by efficiency (%)
Robot controller and teach pendant	72.2	65.8	77.9	0.0	41.0
Safety PLC and interlock circuit	73.1	65.8	79.8	0.1	35.6
AMR fleet management server	73.9	67.6	79.5	0.1	43.1
Machine vision and guidance camera	66.0	59.7	71.7	2.9	70.8
MES and ERP integration gateway	72.0	65.8	77.6	0.0	9.6
Vendor remote access gateway	79.9	72.9	86.0	96.8	100.0

The results in Table 9 support the main findings. Aggregate reduction in the illustrative index averaged 73.6 percent, with 90 percent of draws between 70.7 and 76.3 percent, so the modelled benefit of five moderate layers is stable across the assumed parameter ranges, although it remains a property of the model rather than a measured effect. The vendor remote access gateway ranked first by investment efficiency in 96.8 percent of draws and was in the top three in every draw, which supports brokered vendor access as the most robust first investment within the assumed parameter ranges. The ranking of the middle assets is less stable, with the robot controller, safety controller and mobile robot server each appearing in the top three in roughly 35

to 45 percent of draws. This instability is itself informative: for these assets, the choice of sequence should be guided by the consequence floor and by local evidence rather than by small differences in the efficiency index, consistent with the economic reasoning that investment should follow the value at risk (Gordon and Loeb, 2002).

6.5 Extension to Dependent Layer Failures

The multiplicative residual risk model in Equation (4) assumes that layers fail independently. In practice, layers share common causes: a compromised identity provider can weaken both access control and monitoring, and a poorly governed vendor relationship can undermine segmentation and response at once. Reliability engineering addresses this problem through common cause failure models, of which the beta-factor model is the most widely used (Mosleh et al., 1988; Rausand and Høyland, 2004). Adapting it here, a fraction β of the protective effect is assumed to fail through a shared cause, in which case only the single most effective layer provides protection. Residual risk then becomes Equation (7).

$$R_i^{res,\beta} = R_i^0[(1 - \beta)\prod_{k=1}^5(1 - e_{ik}) + \beta(1 - \max_k e_{ik})] \quad (7)$$

The penalty of dependence relative to the independent case follows from Equation (8), which is always non-negative because the strongest single layer can never remove more risk than all layers together.

$$\Delta R_i^{dep} = \beta R_i^0[(1 - \max_k e_{ik}) - \prod_{k=1}^5(1 - e_{ik})] \quad (8)$$

Table 10 reports the effect of dependence on the illustrative facility for β values of 0, 0.1, 0.2 and 0.3, a range consistent with beta factors commonly applied to redundant protective functions (Rausand and Høyland, 2004).

Table 10. Effect of common cause dependence between layers on residual risk in the illustrative facility

Dependence factor β	Aggregate risk reduction (%)	Total residual risk	Residual risk relative to independence	Vendor gateway reduction (%)	Vision camera reduction (%)
0.0 (independent)	73.7	22.03	1.00	79.9	66.0
0.1	70.2	24.93	1.13	76.9	62.4
0.2	66.8	27.83	1.26	73.9	58.8
0.3	63.3	30.73	1.39	70.9	55.2

As Table 10 shows, a moderate dependence factor of 0.2 raises total residual risk by 26 percent relative to the independent case, and the vision camera, whose strongest single layer is weakest, loses the most protection. Two implications follow. First, the independence assumption is optimistic, so organizations using the model should either estimate β from incident data or report results across a range of β . Second, dependence can be reduced by design: identity, monitoring and recovery functions should not share infrastructure, credentials or administrators, a principle that mirrors the diversity defenses used against common cause failure in safety-instrumented systems (Mosleh et al., 1988).

6.6 Interpretation of the Index, Baselines and an Attack Pathway

Two clarifications govern how the numerical results should be read, in addition to the treatment of ordinal scores in Section 6.7. First, the layer effectiveness values in Table 7 are judgments informed by the control descriptions in the standards and the reviewed studies, not empirical estimates; their role is to show how the framework would be used once an organization substitutes its own estimates. Second, the benefit of the full program must be compared with simpler alternatives. Table 11 reports the aggregate reduction in the illustrative index under four baselines.

Table 11. Reduction in the illustrative risk index under simpler baselines and the full framework (assumed parameters)

Configuration	Layers included	Reduction in aggregate index (%)
Perimeter only	Segmentation	32.2
Governance and segmentation	Layers 1 and 3	38.9
Access control and segmentation	Layers 2 and 3	55.8
Prevention only	Layers 1 to 3	60.2
Full framework	Layers 1 to 5	73.7
Full framework with dependence $\beta = 0.4$	Layers 1 to 5	59.8
Full framework with dependence $\beta = 0.5$	Layers 1 to 5	56.4

As Table 11 shows, under the assumed parameters most of the modelled benefit beyond a perimeter comes from brokered access, and detection and response add a further 13.5 points. With strong dependence between layers, the full framework falls to roughly the level of a prevention-only program, which reinforces the design principle that layers must not share infrastructure. A concrete pathway illustrates the logic. Suppose an attacker compromises a

vendor's support laptop and uses a standing remote connection to reach a robot controller and alter its speed limits. Brokered, time-limited access (Layer 2) removes the standing connection; if access is nonetheless obtained, the conduit rules of Layer 3 confine the session to the cell's engineering interface; Layer 4 flags the parameter change outside a work order window and correlates it with an unexpected force profile; and Layer 5 brings the cell to a safe state through hardwired stops and restores the approved program from an offline golden image. Each step corresponds to a documented weakness in connected robot controllers (Quarta et al., 2017; Pogliani et al., 2019).

6.7 Ordinal Scores, Monotone Transformations and Equal Budgets

Because likelihood and consequence are ordinal ratings, percentage reductions in the index are statements about the selected scoring convention and do not represent ratios of real risk (Cox, 2008). The conclusions that matter for decisions are rankings, and these should not depend on the arbitrary spacing of the 1 to 5 scales. The index was therefore recomputed under three monotone transformations of the scores, a convex mapping of 2 to the power of the score minus one, a squared mapping and a concave logarithmic mapping, and the investment alternatives were compared under an equal budget of 100 cost units, with the safety floor enforced in every allocation method so that the safety programmable logic controller is always funded first. Table 12 reports the results.

Table 12. Rankings under monotone transformations of the ordinal scores and allocations under an equal budget

Analysis	Highest inherent risk	Most efficient first investment	Aggregate index reduction within the convention (%)
Linear scores (base)	Vendor remote access	Vendor remote access	73.7
Convex: $2^{(\text{score} - 1)}$	Vendor remote access	Vendor remote access	74.8
Squared scores	Vendor remote access	Vendor remote access	74.4
Concave: $\log_2(\text{score}) + 1$	Vendor remote access	Vendor remote access	73.3
Budget 100 with safety floor, by efficiency ratio	Safety PLC, vendor access, vision	Not applicable	35.4
Budget 100 with safety floor, by inherent risk	Safety PLC, vendor access, vision	Not applicable	35.4

Analysis	Highest inherent risk	Most efficient first investment	Aggregate index reduction within the convention (%)
Budget 100 with safety floor, safety consequence first	Safety PLC, robot controller	Not applicable	23.6
Budget 100 with safety floor, exact optimum by enumeration	Safety PLC, vision, vendor access	Not applicable	35.4
Budget 100 without safety floor, unconstrained benchmark	Vision, MES gateway, vendor access	Not applicable	35.6

As Table 12 shows, vendor remote access ranks first for both inherent risk and efficiency under every transformation, so this conclusion does not depend on scale spacing, whereas the order of the middle assets changes with the transformation and should not be relied upon. Under an equal budget with the safety floor enforced, allocation by efficiency and by inherent risk both reach the constrained optimum of 35.4 percent found by enumeration, whereas allocation by safety consequence alone yields 23.6 percent because it spends the remaining budget on the costly robot controller. The unconstrained optimum of 35.6 percent omits the safety controller and is reported only as a benchmark; it violates the safety floor and is not an admissible allocation. The floor therefore costs 0.2 points of the illustrative index, a trade-off the framework does not permit to be reversed. Finally, cybersecurity controls must be kept distinct from machinery safety functions. Emergency stops, safety-rated speed and separation monitoring and safety interlocks are designed, validated and assured under functional safety standards independently of cybersecurity (ISO, 2011), and the framework treats them as protected assets whose integrity security must preserve, never as security controls whose effectiveness enters the risk index.

7. Implementation Roadmap

The framework is designed for brownfield environments in which robots and controllers will remain in service for many years, so implementation must proceed in phases that deliver risk reduction early without disrupting production. Table 13 presents a four-phase roadmap. The phases follow the logic of the framework layers, beginning with visibility and governance, then removing the most exposed access paths, then confining potential intrusions, and finally building the detection and response capability that sustains the program. The sequence is consistent with guidance that operational technology security programs should begin with asset identification and architecture before deploying advanced monitoring (Stouffer et al., 2023).

Table 13. Phased implementation roadmap for MLCF-RAS in a brownfield robotic facility

Phase	Indicative duration	Principal activities	Exit criteria
1. Visibility and governance	Months 1 to 3	Asset and connection inventory; zone ownership; risk scoring with consequence weights; secure configuration baselines	Inventory coverage above 95 percent; risk register approved by leadership
2. Access hardening	Months 3 to 6	Brokered vendor and engineering access with multi-factor authentication; removal of default credentials; role-based privileges	All remote sessions brokered and recorded; no shared controller accounts
3. Segmentation	Months 6 to 12	Zone and conduit design; industrial firewalls at cell boundaries; separation of safety networks; wireless isolation for mobile robots	Highest-consequence zones segmented; conduit rules documented and reviewed
4. Detection, response and continuous improvement	Months 9 to 18 and ongoing	OT-aware monitoring; correlation with condition monitoring data; incident playbooks; tabletop exercises; golden images	Mean time to detect measured; two exercises per year; lessons fed into governance

Organizations with constrained budgets can use the efficiency index from Section 6 to decide which zones to address first within each phase, subject to the consequence floor for safety-critical assets (Gordon and Loeb, 2002).

Several practical conditions influence success. Security changes should be scheduled within planned maintenance windows and tested in the same way as other engineering changes, so that they do not cause the unplanned downtime they are meant to prevent. Finally, progress should be communicated in operational terms, such as protected production hours and avoided safety exposure, which resonates with plant leadership more than technical metrics alone (Culot et al., 2019).

8. Discussion

8.1 Theoretical Implications

The framework contributes to the literature on Industry 4.0 cybersecurity in three ways. First, it joins two bodies of work that have developed largely in parallel: technical analyses of robot vulnerabilities and general management frameworks for cybersecurity. By mapping threat

classes from the robotics literature to control layers derived from NIST and IEC guidance, it provides the integrated reference model that reviews by Lezzi et al. (2018) and Culot et al. (2019) found to be missing. The mapping in Table 5 shows that each threat class requires contributions from several layers, which offers a conceptual explanation for why single-control approaches have repeatedly proved insufficient in industrial settings.

Second, the framework treats consequence as multidimensional. Most risk models used in information security express consequence as a single financial loss figure. The consequence-weighted model in Section 6 keeps safety, operational and financial consequences separate until leadership chooses how to combine them, and it adds a consequence floor to prevent safety from being traded against efficiency. This extends the classical scenario-based risk definition of Kaplan and Garrick (1981) to the cyber-physical setting in a way that makes value judgments explicit rather than embedding them silently in loss estimates.

Third, the framework positions maintenance and reliability data as security evidence. The detection layer's use of condition monitoring signals connects two professional communities, reliability engineering and cybersecurity, that rarely share data. This connection is consistent with the observation by Wu and Moon (2017) that cross-domain attacks resemble equipment faults, and it suggests a research agenda in which predictive maintenance models are evaluated not only for their ability to anticipate failures but also for their ability to flag deliberate manipulation.

8.2 Managerial Implications

For managers, the most important implication is that cybersecurity in robotic facilities should be managed as a continuous discipline rather than a project with an end date. The NIST Cybersecurity Framework's emphasis on continuous functions supports this view (NIST, 2018). Organizations should embed security tasks in maintenance procedures, commissioning checklists and change management, and should assign ownership of robotic zones to the engineering leaders who are already accountable for their reliability. Evidence that cybersecurity investment affects financial performance (Ozowara et al., 2022) and models for pricing cyber risk in insurance (Adebayo et al., 2023) support the case for presenting security as a source of measurable value.

A second implication concerns the framing of investment. When cybersecurity is presented only as a cost of avoiding breaches, it competes poorly with production investments. When it is presented as protection of throughput, product quality and worker safety, it becomes part of operational excellence. The efficiency index in Section 6 allows managers to show which investments yield the most protection per dollar, consistent with the economic reasoning of Gordon and Loeb (2002). In the illustrative facility, brokered vendor access emerged as the highest-return investment, a finding that many organizations could act on quickly because it requires little change to the robots themselves.

A third implication concerns people. Technicians and operators are both a source of risk and the most sensitive detection mechanism available, since they notice when machines behave unusually. Parsons et al. (2014) showed that security awareness can be measured and improved through targeted intervention, and Bulgurcu et al. (2010) showed that compliance depends on employees' perceptions of the costs and benefits involved.

8.3 Theoretical Propositions and Research Agenda

The framework implies a set of propositions that can be tested in operating facilities, which moves it from a design artifact toward an empirically falsifiable theory of robotic cybersecurity management. Table 14 states five propositions together with the constructs involved and a research design suited to testing each. The propositions draw on the defense in depth logic of the framework (Mosteiro-Sanchez et al., 2020), on the classical scenario definition of risk (Kaplan and Garrick, 1981) and on evidence that robotic controllers are vulnerable where network trust is assumed (Quarta et al., 2017). Several of the propositions can be tested with records that facilities already hold, such as access logs, configuration audits and incident reports (Cherdantseva et al., 2016).

Table 14. Testable propositions derived from the MLCF-RAS framework

Proposition	Constructs	Suggested test design
P1. Facilities that partition robotic cells into documented zones and conduits experience fewer intrusions that propagate beyond the initial zone than facilities with flat networks	Segmentation maturity; cross-zone propagation events	Multi-site comparison of incident records and red-team exercises
P2. Brokered, time-limited vendor access reduces supply-chain-	Access brokering; compromise via third parties; control cost	Longitudinal incident analysis before and after brokering, with cost

Proposition	Constructs	Suggested test design
originated compromise more per unit cost than any other single control		accounting
P3. Correlating condition-monitoring signals with access and network logs shortens time to detect cyber-physical tampering compared with network-only monitoring	Detection architecture; mean time to detect	Controlled fault-injection and tampering experiments in a testbed cell
P4. Consequence-weighted scoring with a safety floor yields portfolios with lower residual safety exposure than single-loss scoring at equal budget	Scoring method; residual safety-weighted risk	Simulation on real asset registers, validated by expert panels
P5. Embedding configuration verification in preventive maintenance checklists reduces post-maintenance configuration drift	PM procedure design; configuration deviations	Audit sampling before and after procedure change across sites

8.4 Limitations and Future Research

This study has several limitations. The framework was developed through design science and integrative review rather than through field trials, and the demonstration in Section 6 used constructed parameter values. The results therefore show how the model operates but do not estimate the actual risk reduction that any specific organization would achieve. Hevner et al. (2004) emphasize that design artifacts require evaluation in their intended environment, and the next step should be application in operating facilities, with likelihood and effectiveness parameters estimated from incident data, penetration testing and monitoring logs.

The multiplicative residual risk model assumes that layers act independently. Future work could extend the model with dependency structures or simulation methods that capture common-cause failures. Similarly, the effectiveness values in Table 7 were assigned by judgment, and structured elicitation methods or empirical calibration would strengthen them (Cherdantseva et al., 2016).

Finally, the scope of the framework is limited to robotic and autonomous systems in manufacturing, warehousing and distribution. Other sectors, such as healthcare robotics or outdoor autonomous vehicles, face different regulatory regimes and threat profiles. Further research could examine how the layers should be adapted to those contexts and how emerging technologies such as artificial intelligence-based fleet orchestration change the threat landscape.

Yaacoub et al. (2022) observed that robotic security research remains uneven across domains, which suggests considerable scope for comparative studies.

9. Conclusion

Connected robots and autonomous systems have brought cybersecurity onto the factory floor in a form where the consequences of an attack are physical as well as digital. This paper has proposed the Multi-Layered Cybersecurity Framework for Robotics and Autonomous Systems, which organizes protection into five layers of governance and risk assessment, zero trust identity and access, zone and conduit segmentation, operational technology aware detection, and cyber-physical incident response with safe recovery, supported by workforce training, supplier assurance, analytics and measurement.

The framework rests on a threat taxonomy that links spoofing, misconfiguration, supply chain compromise and insecure programming to their physical consequences. Its quantitative model separates safety, operational and financial consequence, computes residual risk across layers and ranks investments by efficiency while protecting safety-critical assets through a consequence floor. Applied to a constructed facility under stated assumptions, the model indicated that layered defenses substantially lower the illustrative risk index and that brokered vendor access is the most robust first investment; field data are needed to establish actual effects.

For leaders of reliability, maintenance and operations, the central message is that cybersecurity belongs alongside preventive maintenance and safety management as a permanent operating discipline. No single control is sufficient, and no program is ever finished, because threats evolve as quickly as the technologies they target. Organizations that build layered defenses, measure their effect and involve the people who run their machines will be better placed to capture the benefits of robotics and automation while keeping their workforce safe and their operations resilient.

Declarations

Funding: This research received no external funding.

Competing interests: The authors declare no competing interests.

Data availability: All parameter values and results are reported in the article. The simulation code, input files, random seeds and distribution settings that reproduce every numerical table are available from the corresponding author on reasonable request.

Ethics approval: Not applicable; the study involved no human participants or personal data.

References

Abomhara, M., & Køien, G. M. (2015). Cyber security and the internet of things: Vulnerabilities, threats, intruders and attacks. *Journal of Cyber Security and Mobility*, 4(1), 65-88. <https://doi.org/10.13052/jcsm2245-1439.414>

Adebayo, A., Adegbite, M. P., & Ahmed, M. O. (2022). Adversarial machine learning in critical infrastructure: A conceptual framework for threat modeling AI enabled OT systems. *World Journal of Innovation and Modern Technology*, 6(1), 184-234. <https://doi.org/10.56201/wjimt.v6.no1.2022.pg184.234>

Adebayo, A., Adepoju, P. A., & Ozowara, D. E. (2023). Conceptual model for cyber risk pricing and insurance structuring in health technology platforms. *Gyanshauryam, International Scientific Refereed Research Journal*, 6(6).

Adegbite, M. P., Adebayo, A., & Ahmed, M. O. (2022). A security architecture model for IT and OT convergence in regulated energy networks: Design principles and governance alignment. *International Journal of Computer Science and Mathematical Theory*, 8(2), 81-132. <https://doi.org/10.56201/ijcsmt.v8.no2.2022.pg81.132>

Ahmed, K. S., & Odejobi, O. D. (2018). Conceptual framework for scalable and secure cloud architectures for enterprise messaging. *IRE Journals*, 2(1), 1-15.

Ahmed, M. O., Adegbite, M. P., & Adebayo, A. (2021). Zero trust architecture for operational technology in North American critical infrastructure: A framework for implementation and resilience optimization. *International Journal of Engineering and Modern Technology*, 7(1), 66-113. <https://doi.org/10.56201/ijemt.vol.7.no1.2021.pg66.113>

Akanbi, O., & Sunday, E. A. (2024). An integrated path-planning and slotting optimization model for AMR-enabled high-density warehousing. *International Journal of Advanced*

Multidisciplinary Research and Studies, 4(6), 3226-3243.
<https://doi.org/10.62225/2583049X.2024.4.6.6181>

Akeju, B., Edivri, J., Ogbole, J. I., Okoruwa, P. O., Fadayomi, O., & Abolaji, T. O. (2018). Conceptual model for insider threat classification and risk modeling in complex digital systems. *Iconic Research and Engineering Journals*, 1(9), 476-492. <https://doi.org/10.64388/IREV119-1713778>

Alcaraz, C., & Zeadally, S. (2015). Critical infrastructure protection: Requirements and challenges for the 21st century. *International Journal of Critical Infrastructure Protection*, 8, 53-66. <https://doi.org/10.1016/j.ijcip.2014.12.002>

Aliliele, C., Mbonu, I. S., & Iwuanyanwu, U. (2023a). A conceptual framework for continuous cloud misconfiguration monitoring and enterprise risk mitigation strategies. *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, 9(10), 373-394. <https://doi.org/10.32628/CSEIT2361071>

Aliliele, C., Mbonu, I. S., & Iwuanyanwu, U. (2023b). A review of API governance and risk prioritization frameworks in modern financial institutions. *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, 9(10), 395-433. <https://doi.org/10.32628/CSEIT2361072>

Ani, U. P. D., He, H., & Tiwari, A. (2017). Review of cybersecurity issues in industrial critical infrastructure: Manufacturing in perspective. *Journal of Cyber Security Technology*, 1(1), 32-74. <https://doi.org/10.1080/23742917.2016.1252211>

Anyanwu, K. C. E. (2019). Developing a maturity model for standardizing IT support operations across geographically distributed multi-office organizations. *Iconic Research and Engineering Journals*, 3(5), 580-598.

Asiedu, W., & Quainoo, R. (2023). Toward maintenance free wireless infrastructure: Simulating performance and reliability in large scale intermittently powered IoT networks. *Gyanshauryam, International Scientific Refereed Research Journal*, 6(1), 489-510.

Borah, S., Aliliele, K. C., Rakshit, S., & Vajjhala, N. R. (2022). Applications of artificial intelligence in software testing. In P. K. Mallick et al. (Eds.), *Cognitive informatics and soft*

computing (Lecture Notes in Networks and Systems, Vol. 375, pp. 727-736). Springer.
https://doi.org/10.1007/978-981-16-8763-1_60

Boyes, H., Hallaq, B., Cunningham, J., & Watson, T. (2018). The industrial internet of things (IIoT): An analysis framework. *Computers in Industry*, 101, 1-12.
<https://doi.org/10.1016/j.compind.2018.04.015>

Bulgurcu, B., Cavusoglu, H., & Benbasat, I. (2010). Information security policy compliance: An empirical study of rationality-based beliefs and information security awareness. *MIS Quarterly*, 34(3), 523-548. <https://doi.org/10.2307/25750690>

Cherdantseva, Y., Burnap, P., Blyth, A., Eden, P., Jones, K., Soulsby, H., & Stoddart, K. (2016). A review of cyber security risk assessment methods for SCADA systems. *Computers & Security*, 56, 1-27. <https://doi.org/10.1016/j.cose.2015.09.009>

Cichonski, P., Millar, T., Grance, T., & Scarfone, K. (2012). *Computer security incident handling guide* (NIST Special Publication 800-61 Revision 2). National Institute of Standards and Technology. <https://doi.org/10.6028/NIST.SP.800-61r2>

Clark, G. W., Doran, M. V., & Andel, T. R. (2017). Cybersecurity issues in robotics. In *2017 IEEE Conference on Cognitive and Computational Aspects of Situation Management (CogSIMA)* (pp. 1-5). IEEE. <https://doi.org/10.1109/COGSIMA.2017.7929597>

Corallo, A., Lazoi, M., & Lezzi, M. (2020). Cybersecurity in the context of Industry 4.0: A structured classification of critical assets and business impacts. *Computers in Industry*, 114, 103165. <https://doi.org/10.1016/j.compind.2019.103165>

Cox, L. A., Jr. (2008). What's wrong with risk matrices? *Risk Analysis*, 28(2), 497-512. <https://doi.org/10.1111/j.1539-6924.2008.01030.x>

Culot, G., Fattori, F., Podrecca, M., & Sartor, M. (2019). Addressing Industry 4.0 cybersecurity challenges. *IEEE Engineering Management Review*, 47(3), 79-86. <https://doi.org/10.1109/EMR.2019.2927559>

- Dosunmu, A. A., & Ogundele, P. O. (2019). Security audit and enterprise risk assessment frameworks for resilient information systems. *Iconic Research and Engineering Journals*, 3(5), 434-447. <https://doi.org/10.64388/IREV3I5-1713225>
- Dosunmu, A. A., & Ogundele, P. O. (2020). Intrusion detection and prevention models for enhancing organizational cyber defense effectiveness. *Iconic Research and Engineering Journals*, 4(6), 310-324. <https://doi.org/10.64388/IREV4I6-1713226>
- Dosunmu, A. A., & Ogundele, P. O. (2021). Incident response and digital forensics strategies for rapid cyber attack containment. *Gyanshauryam, International Scientific Refereed Research Journal*, 4(4), 239-258.
- Dosunmu, A. A., & Ogundele, P. O. (2022). Threat intelligence integration frameworks supporting proactive enterprise cybersecurity decision making. *Gyanshauryam, International Scientific Refereed Research Journal*, 5(3), 397-416.
- Dosunmu, A. A., & Ogundele, P. O. (2023). Cyber threat actor analysis models for strategic enterprise security planning. *Shodhshauryam, International Scientific Refereed Research Journal*, 6(5), 513-531.
- Dosunmu, A. A., & Ogundele, P. O. (2024a). Breach and attack simulation frameworks for continuous validation of enterprise security controls. *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, 10(3), 1100-1119.
- Dosunmu, A. A., & Ogundele, P. O. (2024b). Cyber risk quantification models for prioritizing enterprise security investment decisions. *International Journal of Multidisciplinary Research and Growth Evaluation*, 5(6), 1777-1785. <https://doi.org/10.54660/IJMRGE.2024.5.6.1777-1785>
- Ejimah, A. A., & Aregbe, O. O. (2022). Feasibility study and design of a piezoelectric energy harvesting system for self-powered street lighting on Nigerian high-traffic highways. *Global Journal of Applied, Management and Social Sciences*, 24, 339-347.
- Elhabashy, A. E., Wells, L. J., & Camelio, J. A. (2019). A cyber-physical attack taxonomy for production systems: A quality control perspective. *Journal of Intelligent Manufacturing*, 30(6), 2489-2504.

- Fadayomi, O., Abolaji, T. O., Edivri, J., Ogbole, J. I., Okoruwa, P. O., & Akeju, B. (2019). Risk-based cybersecurity assurance and data availability: Limitations, advances and future research opportunities. *Iconic Research and Engineering Journals*, 2(12), 602-617. <https://doi.org/10.64388/IREV2I12-1713779>
- Gordon, L. A., & Loeb, M. P. (2002). The economics of information security investment. *ACM Transactions on Information and System Security*, 5(4), 438-457. <https://doi.org/10.1145/581271.581274>
- Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design science in information systems research. *MIS Quarterly*, 28(1), 75-105. <https://doi.org/10.2307/25148625>
- Humayed, A., Lin, J., Li, F., & Luo, B. (2017). Cyber-physical systems security: A survey. *IEEE Internet of Things Journal*, 4(6), 1802-1831. <https://doi.org/10.1109/JIOT.2017.2703172>
- Idika, C. N., Salami, E. O., Ijiga, O. M., & Enyejo, L. A. (2021). Deep learning driven malware classification for cloud-native microservices in edge computing architectures. *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, 7(4).
- International Electrotechnical Commission. (2013). *IEC 62443-3-3: Industrial communication networks, network and system security, Part 3-3: System security requirements and security levels*. IEC.
- International Electrotechnical Commission. (2020). *IEC 62443-3-2: Security for industrial automation and control systems, Part 3-2: Security risk assessment for system design*. IEC.
- International Organization for Standardization. (2011). *ISO 10218-1:2011 Robots and robotic devices: Safety requirements for industrial robots, Part 1: Robots*. ISO.
- Jaakkola, E. (2020). Designing conceptual articles: Four approaches. *AMS Review*, 10(1-2), 18-26. <https://doi.org/10.1007/s13162-020-00161-0>
- Jimoh, H. O., Abolle-Okoyeagu, C. J., Ahmed, M. O., & Lawal, N. O. (2023a). Advancing security in IoT-driven critical infrastructure: A focus on smart transportation system. *American Journal of Engineering Research*, 12(12), 33-46.

- Jimoh, H. O., Ahmed, M. O., & Fagbade, M. O. (2023b). The role of frameworks in cybersecurity governance. *Journal of Behavioural Informatics, Digital Humanities and Development Research*, 9(4), 7-16. <https://doi.org/10.22624/AIMS/BHI/V9N4P2>
- Kaplan, S., & Garrick, B. J. (1981). On the quantitative definition of risk. *Risk Analysis*, 1(1), 11-27. <https://doi.org/10.1111/j.1539-6924.1981.tb01350.x>
- Kerns, A. J., Shepard, D. P., Bhatti, J. A., & Humphreys, T. E. (2014). Unmanned aircraft capture and control via GPS spoofing. *Journal of Field Robotics*, 31(4), 617-636. <https://doi.org/10.1002/rob.21513>
- Kolias, C., Kambourakis, G., Stavrou, A., & Voas, J. (2017). DDoS in the IoT: Mirai and other botnets. *Computer*, 50(7), 80-84. <https://doi.org/10.1109/MC.2017.201>
- Kumuyi, O., Akeju, B., Uzoka, E., & Ozowara, D. E. (2023). Framework for blockchain-based cross-border data exchange and regulatory transparency. *International Journal of Multidisciplinary Futuristic Development*, 4(1), 99-113.
- Ladapo, O. O., Dosunmu, A. A., Jooda, D., & Abolaji, T. O. (2018). Lessons learned from offline assessment of security-critical systems: The case of Microsoft Active Directory. *Iconic Research and Engineering Journals*, 2(6), 277-299. <https://doi.org/10.64388/IREV2I6-1717205>
- Ladapo, O. O., Dosunmu, A. A., Jooda, D., & Abolaji, T. O. (2022a). Human-in-the-loop machine learning: A state of the art. *Journal of Frontiers in Multidisciplinary Research*, 3(1), 656-669. <https://doi.org/10.54660/.JFMR.2022.3.1.656-669>
- Ladapo, O. O., Dosunmu, A. A., Jooda, D., & Abolaji, T. O. (2023a). Active Directory attacks steps, types, and signatures. *International Journal of Advanced Multidisciplinary Research and Studies*, 3(6), 2863-2874. <https://doi.org/10.62225/2583049X.2023.3.6.6204>
- Ladapo, O. O., Dosunmu, A. A., Jooda, D., & Abolaji, T. O. (2023b). An IT service management literature review: Challenges, benefits, opportunities, and implementation practices. *International Journal of Advanced Multidisciplinary Research and Studies*, 3(6), 2875-2889. <https://doi.org/10.62225/2583049X.2023.3.6.6205>

- Ladapo, O. O., Jooda, D., Dosunmu, A. A., & Abolaji, T. O. (2019). Implementation of Active Directory for efficient management of enterprise networks. *Iconic Research and Engineering Journals*, 3(4), 608-627. <https://doi.org/10.64388/IREV3I4-1717206>
- Ladapo, O. O., Jooda, D., Dosunmu, A. A., & Abolaji, T. O. (2022b). Navigating digital transformation: Best practices for cloud migration strategies in the enterprise. *Journal of Frontiers in Multidisciplinary Research*, 3(1), 643-655. <https://doi.org/10.54660/JFMR.2022.3.1.643-655>
- Ladapo, O. O., Jooda, D., Dosunmu, A. A., & Abolaji, T. O. (2023c). Systematic literature review on security access control policies and techniques based on privacy requirements in a BYOD environment. *International Journal of Advanced Multidisciplinary Research and Studies*, 3(6), 2890-2904. <https://doi.org/10.62225/2583049X.2023.3.6.6206>
- Langner, R. (2011). Stuxnet: Dissecting a cyberwarfare weapon. *IEEE Security & Privacy*, 9(3), 49-51. <https://doi.org/10.1109/MSP.2011.67>
- Lasi, H., Fettke, P., Kemper, H.-G., Feld, T., & Hoffmann, M. (2014). Industry 4.0. *Business & Information Systems Engineering*, 6(4), 239-242. <https://doi.org/10.1007/s12599-014-0334-4>
- Lee, J., Bagheri, B., & Kao, H.-A. (2015). A cyber-physical systems architecture for Industry 4.0-based manufacturing systems. *Manufacturing Letters*, 3, 18-23. <https://doi.org/10.1016/j.mfglet.2014.12.001>
- Lezzi, M., Lazoi, M., & Corallo, A. (2018). Cybersecurity for Industry 4.0 in the current literature: A reference framework. *Computers in Industry*, 103, 97-110. <https://doi.org/10.1016/j.compind.2018.09.004>
- Mbonu, I. S., Aliliele, C., Iwuanyanwu, U., & Oluoha, O. M. (2018). A conceptual framework for legal and ethical risk modeling in enterprise data protection governance systems. *Iconic Research and Engineering Journals*, 2(2), 207-226. <https://doi.org/10.64388/IREV2I2-1714911>
- Mbonu, I. S., Aliliele, C., Iwuanyanwu, U., & Uzoka, E. (2020a). A review of identity and access management integration strategies in hybrid and multi cloud environments. *International Journal of Multidisciplinary Research and Growth Evaluation*, 1(5), 795-810. <https://doi.org/10.54660/IJMRGE.2020.1.5.795-810>

Mbonu, I. S., Aliliele, C., Iwuanyanwu, U., & Uzoka, E. (2021). Advances in artificial intelligence techniques for secure software testing and automated regression control mechanisms. *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, 7(5), 468-496. <https://doi.org/10.32628/CSEIT217565>

Mbonu, I. S., Aliliele, C., Iwuanyanwu, U., & Uzoka, E. (2022a). A conceptual framework for AI enabled IT general controls and SOX audit automation processes. *Gyanshauryam, International Scientific Refereed Research Journal*, 5(5), 384-414. <https://doi.org/10.32628/GISRRJ2256239>

Mbonu, I. S., Aliliele, C., Uzoka, E., & Oluoha, O. M. (2019a). A review of comparative data protection regulations and secure cloud implementation strategies across jurisdictions. *Iconic Research and Engineering Journals*, 2(9), 482-501. <https://doi.org/10.64388/IREV2I9-1714912>

Mbonu, I. S., Iwuanyanwu, U., Aliliele, C., & Uzoka, E. (2020b). A conceptual framework for agile supply chain digital transformation with embedded IT risk and ISO compliance controls. *Iconic Research and Engineering Journals*, 3(11), 566-593. <https://doi.org/10.64388/IREV3I11-1714916>

Mbonu, I. S., Iwuanyanwu, U., Aliliele, C., & Uzoka, E. (2020c). Advances in infrastructure as code governance for secure terraform based enterprise cloud deployments. *International Journal of Multidisciplinary Research and Growth Evaluation*, 1(5), 811-828. <https://doi.org/10.54660/IJMRGE.2020.1.5.811-828>

Mbonu, I. S., Iwuanyanwu, U., Aliliele, C., & Uzoka, E. (2022b). A review of data protection impact assessment models in multi cloud financial infrastructure systems. *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, 8(1), 589-623. <https://doi.org/10.32628/CSEIT25442>

Mbonu, I. S., Iwuanyanwu, U., Aliliele, C., & Uzoka, E. (2022c). Advances in cloud identity and access governance optimization in large scale AWS enterprise environments. *Shodhshauryam, International Scientific Refereed Research Journal*, 5(3), 403-438. <https://doi.org/10.32628/SHISRRJ225490>

- Mbonu, I. S., Iwuanyanwu, U., Uzoka, E., & Oluoha, O. M. (2019b). Advances in enterprise log analytics and automated incident response architectures using Python and SIEM platforms. *Iconic Research and Engineering Journals*, 3(2), 1000-1019. <https://doi.org/10.64388/IREV3I2-1714915>
- Mitchell, R., & Chen, I.-R. (2014). A survey of intrusion detection techniques for cyber-physical systems. *ACM Computing Surveys*, 46(4), Article 55. <https://doi.org/10.1145/2542049>
- Mosleh, A., Fleming, K. N., Parry, G. W., Paula, H. M., Worledge, D. H., & Rasmuson, D. M. (1988). *Procedures for treating common cause failures in safety and reliability studies: Procedural framework and examples* (NUREG/CR-4780, Vol. 1). U.S. Nuclear Regulatory Commission.
- Mosteiro-Sanchez, A., Barcelo, M., Astorga, J., & Urbieto, A. (2020). Securing IIoT using defence-in-depth: Towards an end-to-end secure Industry 4.0. *Journal of Manufacturing Systems*, 57, 367-378. <https://doi.org/10.1016/j.jmsy.2020.10.011>
- National Institute of Standards and Technology. (2018). *Framework for improving critical infrastructure cybersecurity, Version 1.1*. NIST. <https://doi.org/10.6028/NIST.CSWP.04162018>
- National Institute of Standards and Technology. (2024). *The NIST Cybersecurity Framework (CSF) 2.0* (NIST CSWP 29). U.S. Department of Commerce. <https://doi.org/10.6028/NIST.CSWP.29>
- Ogbole, J. I., Okoruwa, P. O., Fadayomi, O., Abolaji, T. O., Edivri, J., & Akeju, B. (2021). Conceptual model for identity-centric zero trust architecture in enterprise security governance. *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, 7(5), 393-415. <https://doi.org/10.32628/IJSRCSEIT217562>
- Ohm, M., Plate, H., Sykosch, A., & Meier, M. (2020). Backstabber's knife collection: A review of open source software supply chain attacks. In C. Maurice, L. Bilge, G. Stringhini, & N. Neves (Eds.), *Detection of Intrusions and Malware, and Vulnerability Assessment (DIMVA 2020)* (Lecture Notes in Computer Science, Vol. 12223, pp. 23-43). Springer. https://doi.org/10.1007/978-3-030-52683-2_2

Ojukwu, J., Oyesiji, S. O., & Nwakamma, S. (2023). Cyber-physical ransomware defense in operational technology using physics-informed detection, digital twins, network telemetry, and resilient recovery. *International Journal of Computer Science and Mathematical Theory*, 9(5), 193-228. <https://doi.org/10.56201/ijcsmt.v9.no5.2023.pg193.228>

Okoruwa, P. O., Fadayomi, O., Akeju, B., Edivri, J., Ogbole, J. I., & Abolaji, T. O. (2020). Conceptual model for privacy-centric security engineering in digital and cloud computing systems. *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, 7(5), 371-389.

Onche, V. O., & Ogonna, C. S. (2022). Password practices, authentication behaviour, and data-protection compliance among secretarial professionals in higher education. *International Journal of Multidisciplinary Research and Growth Evaluation*, 3(6), 1001-1015. <https://doi.org/10.54660/IJMRGE.2022.3.6.1001-1015>

Onche, V. O., Adegbite, M. P., & Ogonna, C. S. (2023). BYOD adoption and information-security risk in university administrative offices: Evidence from Nigerian federal universities. *International Journal of Multidisciplinary Research and Growth Evaluation*, 4(6), 1549-1563. <https://doi.org/10.54660/IJMRGE.2023.4.6.1549-1563>

Onche, V. O., Ogonna, C. S., & Adegbite, M. P. (2020). Information-security awareness and records confidentiality in Nigerian university registries: A behavioural assessment. *International Journal of Multidisciplinary Research and Growth Evaluation*, 1(5), 1032-1044. <https://doi.org/10.54660/IJMRGE.2020.1.5.1032-1044>

Oshoba, T. O., Hammed, N. I., & Odejebi, O. D. (2019). Secure identity and access management model for distributed and federated systems. *IRE Journals*, 3(4), 550-567.

Oshoba, T. O., Hammed, N. I., & Odejebi, O. D. (2020). Blockchain-enabled compliance and audit trail model for cloud configuration management. *Journal of Frontiers in Multidisciplinary Research*, 1(1), 193-201.

Ozowara, D. E., Adebayo, A., & Anunagba, C. O. (2022). A systematic review of cybersecurity investments and their impact on healthcare financial performance. *Shodhshauryam, International Scientific Refereed Research Journal*, 5(1), 404-427.

- Parsons, K., McCormac, A., Butavicius, M., Pattinson, M., & Jerram, C. (2014). Determining employee awareness using the Human Aspects of Information Security Questionnaire (HAIS-Q). *Computers & Security*, 42, 165-176. <https://doi.org/10.1016/j.cose.2013.12.003>
- Peffer, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2007). A design science research methodology for information systems research. *Journal of Management Information Systems*, 24(3), 45-77. <https://doi.org/10.2753/MIS0742-1222240302>
- Pliatsios, D., Sarigiannidis, P., Lagkas, T., & Sarigiannidis, A. G. (2020). A survey on SCADA systems: Secure protocols, incidents, threats and tactics. *IEEE Communications Surveys & Tutorials*, 22(3), 1942-1976. <https://doi.org/10.1109/COMST.2020.2987688>
- Pogliani, M., Quarta, D., Polino, M., Vittone, M., Maggi, F., & Zanero, S. (2019). Security of controlled manufacturing systems in the connected factory: The case of industrial robots. *Journal of Computer Virology and Hacking Techniques*, 15(3), 161-175. <https://doi.org/10.1007/s11416-019-00329-8>
- Quarta, D., Poglian, M., Polino, M., Maggi, F., Zanchettin, A. M., & Zanero, S. (2017). An experimental security analysis of an industrial robot controller. In *2017 IEEE Symposium on Security and Privacy (SP)* (pp. 268-286). IEEE. <https://doi.org/10.1109/SP.2017.20>
- Rausand, M., & Høyland, A. (2004). *System reliability theory: Models, statistical methods, and applications* (2nd ed.). John Wiley & Sons.
- Reason, J. (1997). *Managing the risks of organizational accidents*. Ashgate.
- Rose, S., Borchert, O., Mitchell, S., & Connelly, S. (2020). *Zero trust architecture* (NIST Special Publication 800-207). National Institute of Standards and Technology. <https://doi.org/10.6028/NIST.SP.800-207>
- Saltzer, J. H., & Schroeder, M. D. (1975). The protection of information in computer systems. *Proceedings of the IEEE*, 63(9), 1278-1308. <https://doi.org/10.1109/PROC.1975.9939>
- Shittu, H., & Nabil, M. (2023). Smart supply chain management with attribute-based encryption access control. In *2023 IEEE 13th Annual Computing and Communication Workshop and Conference (CCWC)* (pp. 198-204). IEEE. <https://doi.org/10.1109/CCWC57344.2023.10099268>

Shittu, H., Olagunju, F., & Shittu, M. A. (2024). Cyber physical resilience in digital substations: IoT enabled adaptive protection for secure DER integration. *International Journal of Science, Architecture, Technology and Environment*, 1(3), 81-99. <https://doi.org/10.63680/ijstate032532.08>

Shittu, M. A., Adeniji, I. O., Shittu, H., & Opara, I. S. (2022). Blockchain-assisted secure data exchange architectures for SCADA-controlled power systems. *Iconic Research and Engineering Journals*, 6(3), 359-379. <https://doi.org/10.64388/IREV6I3-1714041>

Shittu, M. A., Shittu, H. A., Adeleke, O. J., & Adedokun, O. J. (2023). Digital twin modeling for real time monitoring and fault detection in smart substations. *International Journal of Industrial Engineering Research and Development*, 14(2), 25-44. https://doi.org/10.34218/IJIERD_14_02_003

Sisinni, E., Saifullah, A., Han, S., Jennehag, U., & Gidlund, M. (2018). Industrial internet of things: Challenges, opportunities, and directions. *IEEE Transactions on Industrial Informatics*, 14(11), 4724-4734. <https://doi.org/10.1109/TII.2018.2852491>

Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. *Journal of Business Research*, 104, 333-339. <https://doi.org/10.1016/j.jbusres.2019.07.039>

Stouffer, K., Pease, M., Tang, C., Zimmerman, T., Pillitteri, V., Lightman, S., Hahn, A., Saravia, S., Sherule, A., & Thompson, M. (2023). *Guide to operational technology (OT) security* (NIST Special Publication 800-82 Revision 3). National Institute of Standards and Technology. <https://doi.org/10.6028/NIST.SP.800-82r3>

Syed, N. F., Shah, S. W., Shaghaghi, A., Anwar, A., Baig, Z., & Doss, R. (2022). Zero trust architecture (ZTA): A comprehensive survey. *IEEE Access*, 10, 57143-57179. <https://doi.org/10.1109/ACCESS.2022.3174679>

Torraco, R. J. (2005). Writing integrative literature reviews: Guidelines and examples. *Human Resource Development Review*, 4(3), 356-367. <https://doi.org/10.1177/1534484305278283>

Tuptuk, N., & Hailes, S. (2018). Security of smart manufacturing systems. *Journal of Manufacturing Systems*, 47, 93-106. <https://doi.org/10.1016/j.jmsy.2018.04.007>

Villani, V., Pini, F., Leali, F., & Secchi, C. (2018). Survey on human-robot collaboration in industrial settings: Safety, intuitive interfaces and applications. *Mechatronics*, 55, 248-266. <https://doi.org/10.1016/j.mechatronics.2018.02.009>

Webster, J., & Watson, R. T. (2002). Analyzing the past to prepare for the future: Writing a literature review. *MIS Quarterly*, 26(2), xiii-xxiii.

Wu, M., & Moon, Y. B. (2017). Taxonomy of cross-domain attacks on cybermanufacturing system. *Procedia Computer Science*, 114, 367-374.

Xu, L. D., Xu, E. L., & Li, L. (2018). Industry 4.0: State of the art and future trends. *International Journal of Production Research*, 56(8), 2941-2962. <https://doi.org/10.1080/00207543.2018.1444806>

Yaacoub, J.-P. A., Noura, H. N., Salman, O., & Chehab, A. (2022). Robotics cyber security: Vulnerabilities, attacks, countermeasures, and recommendations. *International Journal of Information Security*, 21(1), 115-158. <https://doi.org/10.1007/s10207-021-00545-8>