Journal of Computer Science and Technology Studies
ISSN: 2709-104X
DOI: 10.32996/jcsts
Journal Homepage: www.al-kindipublisher.com/index.php/jcsts

JCSTS
AL-KINDI CENTER FOR RESEARCH AND DEVELOPMENT

| RESEARCH ARTICLE

A Framework for Securing Agentic AI Workflows and Quantum-Resistant Communication in U.S. Critical Infrastructure Networks

Md Humayun Kabir[1], Md Al Mamun Siddike[2], Md Riyad Uddin[3], and MD RAZIB[4]

[1] *Master of science in information technology, Westcliff University*

[2] *MS IN BUSINESS ANALYTICS, Trine University*

[34]*Westcliff University*

Corresponding Author: Md Humayun Kabir, **E-mail**: Humayun9152@gmail.com

| ABSTRACT

For artificial intelligence (AI) systems designed to learn and act without human intervention, the events of this rapid integration into U.S. critical infrastructure networks and the rapid pace of approach to cryptographically relevant quantum computers (CRQCs) create a dual convergence of threat that is not adequately met by current frameworks. Agentic AI systems have additional attack surfaces, such as prompt injection, non-human identity (NHI) exploitation, and cascading multi-agent failures, due to their ability to operate autonomously, take multiple steps, and coordinate with other agents. At the same time, existing public-key cryptographic standards, such as RSA and elliptic curve cryptography (ECC), are under attack with harvest-now-decrypt-later (HNDL) attacks to hack encrypted communications in operational technology (OT) and industrial control system (ICS) environments. This article suggests security in both threat domains by leveraging three layers of security: Layer 1: Security of the agentic AI workflow through identity management, prompt sanitization and behavior monitoring; Layer 2: Quantum resistant communication with migration to NIST standardized post-quantum cryptographic (PQC) algorithms (FIPS 203, 204, 205); and Layer 3: Governance and compliance integration with CISA, NERC CIP and the NIST AI Risk Management Framework. Compliance is verified with a cross-walk analysis to current standards, and the framework is tested in scenarios of representative critical infrastructure in the energy, water, and financial sectors. Results suggest that this framework offers a roadmap for RTOs to use in preparing for the deployment of agentic AI and the migration to post-quantum cryptographic algorithms.

| KEYWORDS

Agentic AI, post-quantum cryptography, critical infrastructure security, zero trust architecture, multi-agent systems, NIST PQC standards, prompt injection, quantum-resistant communication

| ARTICLE INFORMATION

ACCEPTED: 15 April 2026 **PUBLISHED:** 10 May 2026 **DOI:** 10.32996/jcsts.2026.8.6.12

1. Introduction

The United States has some of the most complex and interdependent systems of critical infrastructures in the world, including energy infrastructure, water treatment plants, financial systems, transportation and communications networks, and health care systems. They are the backbone for the security, economic stability, and well-being of the nation. Two phenomena have cast profound changes to the cybersecurity reality of these networks over last ten years: Agentic Artificial Intelligence and the maturation of Quantum Computing.

Agentic AI (autonomous execution and coordination of multi-step tasks with the ability to call tools/APIs, adapt to feedback) has ceased to be Theory. With tools like LangGraph, AutoGen and CrewAI facilitating enterprise deployment, industry analysts at Gartner have forecast that by 2028, more than a third of enterprise applications will be using agentic AI (Kaur et al., 2025). In

critical infrastructure environments, AI agents are increasingly being integrated into security operations centers (SOCs), anomaly detection systems, grid management systems, and incident response automation. As great as these capabilities are to the operation they are also a markedly new attack surface one that away from the traditional cybersecurity paradigm.

Unlike conventional software weaknesses, AI agents can be used against them in ways that take advantage of the semantic and autonomous nature which makes them valuable. It has been reported of prompt injection attacks, where malicious instructions inserted in external data sources trick agents into running unauthorized code (McHugh et al., 2025; Tanaka & Okonkwo, 2025), have proven resilient to traditional input filters. Other threat vectors exist that don't have a one-to-one equivalent in pre-AI security models, such as non-human identity (NHI) exploitation, lateral movement through agent privileges, or cascading failures in multi-agent orchestration pipelines (Hammond et al., 2025; Li et al., 2025).

At the same time, quantum computing has gone from theoretical danger to imminent operational threat. A quantum resistant encryption is a term that accumulated a lot of attention because of NIST standards set in May 2022 (NIST NSS 10), which includes a recommendation for government agencies with moderate-to-critical systems to start moving to quantum resistant encryption before 2035. This migration was finalized in August 2024 with three PQD cryptography standards from the National Institute of Standards and Technology (NIST) published by the U.S. Department of Commerce (NIST, 2024a, 2024b, 2024c). This promotes the harvest-now-decrypt-later (HNDL) attack paradigm (Chen & Liu, 2025) in which the attacker might be able to broadcast, store, and decrypt the signals afterwards with quantum technology. The need to make this transition is enhanced by the harvest-now-decrypt-later (HNDL) attack model, where an attacker can capture and store encrypted messages today, and decrypt them later when quantum technology becomes feasible, (Chen & Liu, 2025). This threat exists and is active for critical infrastructure where all communications are of value for some period of time such as national security measurements or operational significance.

Although related to both these threats, the current frameworks are singular in their approach. The NIST Cybersecurity Framework (CSF 2.0), the AI Risk Management Framework (AI RMF 1.0), as well as Zero Trust Architecture guidance (SP 800-207), all help but cover parts only. All lack a combined model of operationally deployable agentic AI workflow security and post-quantum communication resilience for use in U.S. critical infrastructure.

This article fills that void. Our main offering is a three-layered, integrated security approach Agentic-AI and Quantum-Resistant Infrastructure Framework (AQIF) that offers an overlay of guidance for both securing agentic AI deployments and transitioning communications infrastructure to quantum-resistant alternatives.

Table 1: Comparison of Existing Cybersecurity Frameworks and Coverage Gaps Relevant to Agentic AI and Post-Quantum Communication in Critical Infrastructure

Framework	Agentic AI Security	PQC Migration	OT/ICS Guidance	NHI Management	AIBOM Requirements
NIST CSF 2.0	Partial	Not Addressed	Partial	Not Addressed	Not Addressed
NIST AI RMF 1.0	Partial	Not Addressed	Not Addressed	Not Addressed	Not Addressed
Zero Trust SP 800-207	Partial	Not Addressed	Partial	Partial	Not Addressed
NIST FIPS 203/204/205	Not Addressed	Full	Not Addressed	Not Addressed	Not Addressed
CISA Secure by Design	Partial	Partial	Partial	Not Addressed	Not Addressed
AQIF (Proposed)	**Full**	**Full**	**Full**	**Full**	**Full**

Note. Coverage assessed against twelve AQIF control dimensions. Full = comprehensive guidance provided; Partial = some relevant guidance; = not addressed. AQIF = Agentic-AI and Quantum-Resistant Infrastructure Framework; CSF = Cybersecurity Framework; AI RMF = AI Risk Management Framework; ZTA = Zero Trust Architecture; CISA = Cybersecurity and Infrastructure Security Agency; PQC = post-quantum cryptography; OT/ICS = operational technology/industrial control systems; NHI = non-human identity; AIBOM = AI Bill of Materials.

2. Background and Related Work

2.1 Agentic AI in Critical Infrastructure

AI in critical infrastructure has taken a path from simply following rules to becoming more independent and goal-oriented. In initial AI deployments, the AI system was used to predict future maintenance needs, anticipate customer demand, or pinpoint anomalies, all of which required human intervention to take action. The present agentic AIs are a qualitatively different system, however. These systems include planning loops, persistent memory, and the ability to activate or coordinate with other agents, and therefore can be regarded as autonomous digital workers that live inside working environments.

The implications to security of this change are significant. As an AI's recommendation moves to action as querying a database, issuing a control command, or calling an external service the impacts of a potential security breach increase. In critical infrastructure, operational technology (OT) systems govern physical processes like power generation, water treatment, and controlling the pressure in pipelines, making the combination of agentic AI and safety-critical control loops present unique risk scenarios that current cybersecurity frameworks are not yet fully defined (Paulraj et al., 2025).

Agentic architectures have a number of threat vectors peculiar to this architecture. Prompt injection, both direct (when attacks happen in user input to AI systems) and indirect (when adversarial content is injected into data accessed from outside resources), is the most well-documented agentic AI vulnerability (Tanaka & Okonkwo, 2025; McHugh et al., 2025). The processing of natural language instructions and data in agentic systems is usually not separated by hard syntactic rules, which makes it easy to inject instructions that can displace legitimate instructions that are not easily recognizable or preventable by traditional input checking. Lee et al. (2024) showed that the propagation of prompt is possible in multi-agent systems – which they called "prompt infection" – when a compromised agent injects malicious prompt into its message to other agents, and thus malicious prompt can snowball down a whole orchestration pipeline.

A second major category of agentic threats is called non-human identity (NHI) exploitation. AI agents need long-lived access credentials (API Keys, Service Accounts, Authentication Tokens) to communicate with external systems. Cntrls are often issued using general, long-lasting access permissions and are outside of the identity management systems associated with humans (Hammond et al., 2025). Lennon et al. (2025) highlight the importance of NHI management in the trust, risk, and security management (TRiSM) of agentic multi-agent systems: With the rapid growth of machine identities in enterprise environments, there are emerging gaps in security that are actively exploited by adversaries.

In multi-agent coordination, there are added risks that exist outside of single-agent deployments. Hammond et al. (2025) categorize multi-agent failure modes in a structured taxonomy, with miscoordination, conflict and collusion being three key failure modes. Shah and Wang (2025) further works to determine specific security issues in multi-agent systems, such as the challenge of undertaking authorization boundaries when propagation takes place between agents, and the danger of emergent behaviors that may result from agent interaction which was not expected by the system designer.

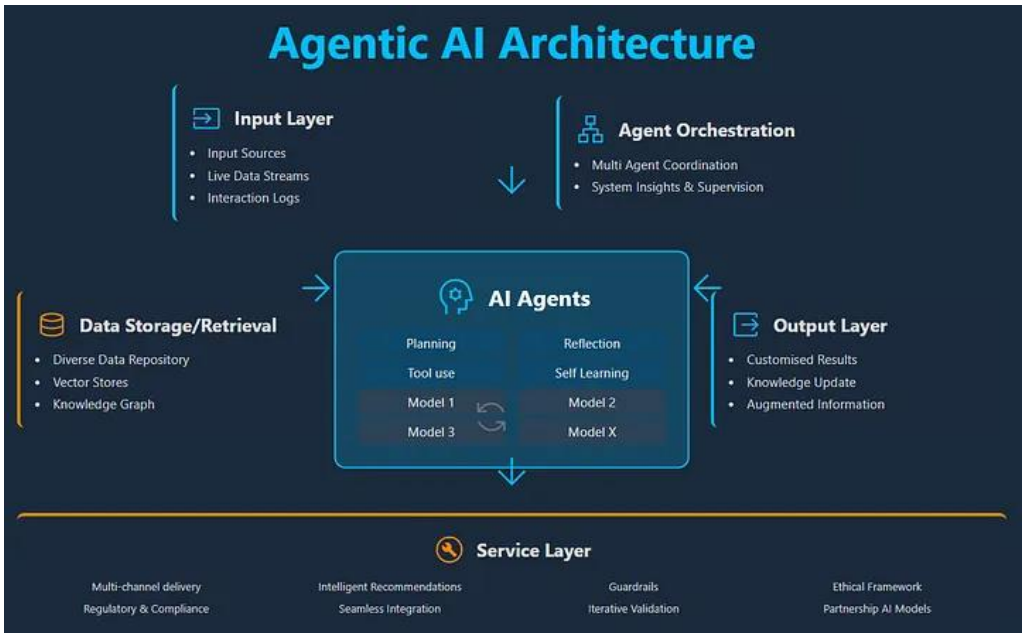

Figure 1. Architecture of an Agentic AI System Deployed in Critical Infrastructure

2.2 Quantum Computing Threats to Communication Security

The potential of quantum computing to compromise traditional cryptographic systems is well known in the literature, and has advanced from theory to an operating planning horizon. Two of the best-known public-key cryptographic systems used in current network infrastructure are the RSA and elliptic curve cryptography (ECC) the latter was formulated in 1994 by Shor's algorithm which provides a polynomial-time solution to the integer factorization and discrete logarithm problems on which these systems are based. One particular quantum computer (cryptographically relevant quantum computer or CRQC) would be capable of breaking an RSA-2048 encryption key in hours instead of billions of years as done by a classical machine.

The harvest-now-decrypt-later (HNDL) attack model makes this impending weakness into a real threat. It is thought to be especially nation-state adversarial actors with the most advanced technical abilities who are systematically capturing and storing encrypted communications with the aim of decrypting them upon the realization of CRQC capabilities (Rahman et al., 2025; Chen & Liu, 2025). HNDL can pose an active and ongoing threat in critical infrastructure applications even if the actual time of CRQC is a long way off, as such communications have long-term operational or national security importance, such as grid control signals, parameters of water treatment, financial transaction records, or classified government communications.

The critical need for post-quantum cryptographic migration is compounded by the difficulty of cryptographic infrastructure migrations for operational environments in OT/ICS. In contrast to IT systems, OT and ICS components are typically deployed for 15-25 years, have limited computing power and have strict availability requirements that restrict maintenance periods. The cryptographic agility of key pairs, which is also known as the rekeying of the cryptographic keys, provoked such complexity and time at the protocol level that protocols that governed the industrial control communications such as DNP3, Modbus, and IEC 61850 were making the migration of PQC technologically challenging and time consuming (Wang et al., 2025; Hassan et al., 2025).

Table 2: Quantum Computing Threat Timeline and Estimated Risk Exposure by Critical Infrastructure Sector

Infrastructure Sector	Primary Cryptographic Exposure	HNDL Risk Level	Estimated Migration Urgency	Recommended Action Timeline
Energy (Grid/SCADA)	TLS, DNP3, IEC 61850 PKI	Critical	Immediate	Begin inventory Q1 2025; hybrid PQC by 2027
Water & Wastewater	Modbus/TCP, remote access VPNs	High	High	Inventory complete by 2026; pilot PQC by 2028
Financial Services	TLS 1.3, HSM key storage, SWIFT	Critical	Immediate	Hybrid TLS PQC deployed by end of 2026
Healthcare	HL7 FHIR APIs, DICOM, EHR encryption	High	High	AIBOM + PQC roadmap by 2026; deploy by 2029
Transportation	GPS signal auth, V2X communications	Medium	Moderate	Standards development phase; deploy by 2030
Communications	BGP security, internet PKI, VoIP	Critical	Immediate	RPKI + PQC migration underway; complete by 2027

Note. HNDL = harvest-now-decrypt-later. Risk levels reflect current exposure from HNDL attacks and proximity of CRQC timelines to system operational lifespans. Migration urgency: Immediate = action required 2025–2026; High = action required by 2027; Moderate = standards development phase. CRQC = cryptographically relevant quantum computer; PKI = public key infrastructure; SCADA = supervisory control and data acquisition.

2.3 Existing Frameworks and Their Limitations

There are a few authoritative frameworks in part that give guidance regarding the two dangers of concern in this article. The NIST Cybersecurity Framework (CSF) 2.0 (2024) has six functions Govern; Identify; Protect; Detect; Respond; Recover and

explicitly introduces supply chain risk management and governance considerations. The challenges of agentic AI, however, are not attended to by CSF 2.0, which doesn't offer specifics on PQC migration strategies or threat vectors specific to GHPs.

Four functions Map, Measure, Manage, and Govern are included in the NIST AI Risk Management structure (AI RMF 1.0), which creates a governance structure to handle AI-related risks during the course of the AI lifecycle. The AI RMF does not specifically address the operational security vulnerabilities associated with agentic deployments of AI in critical scenarios, despite including features of "trustworthiness," bias, and explainability. One coming development to patch this gap involves NIST's December 2025 preliminary draft of the Cybersecurity Framework Profile for Artificial Intelligence or AI Framework Profile.

At its core, Zero Trust Architecture (ZTA), as embodied in NIST SP 800-207, will have no implicit trust based on network location or prior context for any given human or machine entity. The principles of ZTA are applicable in agentic AI environments, where agents might need to interact with other agents or organizations, and continuous fine-grained authorizations are needed for each action. Jain and Singh (2025) consider that the most universally adopted ZTA components are authentication, authorization, and access control; others are underdeveloped and include auditing, orchestration, and environmental perception. Alladi et al. (2025) illustrate how ZTA principles can be put into practice in industrial network security, integrating zero-trust controls with AI-driven anomaly detection within ICS networks.

All of these frameworks encompass crucial guiding principles, but a key missing element of each is that agentic AI workflow security does not meet post-quantum communication resilience in a single framework tailored to the U.S. critical infrastructure landscape and use cases. This coverage gap is summarised systemically in Table 1.

3. Threat Landscape

3.1 Agentic AI Attack vectors

In the world of critical infrastructure, agentic AI creates a new threat surface with four key categories of attack vectors, each qualitatively different from pre-AI threats.

These attacks are prompt injection and prompt manipulation attacks, which capitalises on the core architectural characteristic of LLM-based agents: the failure to reliably identify legitimate instructions from adversarial data. Direct prompt injection is a way to inject malicious instructions into an input directly fed to the agent as the users' instructions, while indirect prompt injection is a way to inject adversarial content into an external data source such as emails, documents, web pages, or database records that are retrieved and processed automatically by the agent (Tanaka & Okonkwo, 2025). Often, in the critical infrastructure use case, sensor data, log files, and reports are regularly fed (upstream of the agent's processing pipeline) to agents, making indirect prompt injection particularly harmful. McHugh et al. (2025) provide a timeline of the progress towards hybrid prompt injection attacks, where adversarial AI techniques are combined with other traditional cybersecurity attack methods like cross-site scripting and SQL injection to form attack chains where AI safety filters are circumvented and traditional security solutions are bypassed.

Tool-call hijacking and the ability to make unauthorized calls to APIs is a second big vector. The value of agentic systems lies in how they enable and allow you to call out and invoke external tools and APIs which can be used to issue commands to systems, query databases, initiate automated processes. If an agent gets compromised, for example, by prompt injection in an application or the theft of its credentials, the tool keys of the agent are the execution path. Deploying agents is also a common occurrence in industrial settings where they may be granted access to and interaction with SCADA systems, programmable logic controllers (PLCs) or distributed control systems (DCS) and potentially cause physical disruption of processes, with ramifications that can ripple through other sectors of infrastructure that is interconnected.

Agent impersonation and identity spoofing are due to limited identity system that currently regulates communications between agents. Many times the agents used in the implementation of multi-agent systems are asked to follow instructions from other peer agents without a proper verification of the identification in some specific way. An attacker who compromises an upstream agent in a multi-agent pipeline or who is able to spoof messages as if they came from a trusted orchestrator can manipulate downstream agents without even raising the more typical security alarms (Shah & Wang, 2025). This is especially threatening in distributed multi-agent architectures where the agents are not in the same organization.

Lateral movement with AI agent permissions is leveraging the wide and enduring permissions that are usually afforded to AI agents. Agents need to span security domains as they need to interact with multiple systems in order to perform their operation. Once an agent is compromised and has global reach, it can systematically move from one of these areas into the other, stealing information or hiding and setting up payloads in a systematic movement of attackers between connected OT and IT networks, which may go undetected by network-based intrusion detection systems (IDS) tuned to look for lateral movement at human speeds.

Threat Model Summary:

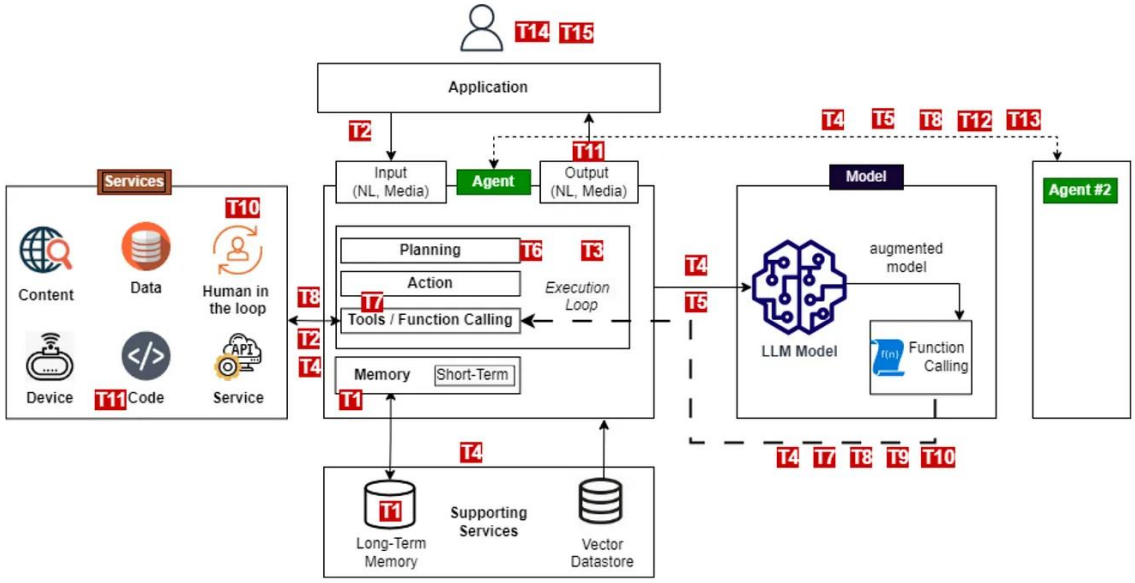

Figure 2. Threat Surface Map for Agentic AI Workflows in Critical Infrastructure Environments. Threat identifiers (T1–T15) indicate documented attack vectors across agent memory, orchestration, tool invocation, external services, model interaction, and inter-agent communication pathways.

3.2 Quantum-Era Communication Vulnerabilities

Quantum threat to critical infrastructure communications is where long-lived, high-value data are protected with public-key cryptography or they control safety critical processes. For structural reasons, operational technology environments are disproportionately exposed. Technically, the substitution of algorithms is difficult in many OT systems because they often use legacy protocols (such as DNP3, Modbus, IEC 61850, or BACnet which are not inherently flexible about algorithms. Second, because most of the OT systems installed in the past also can, and are expected to, remain in operation long after the approximate time of a CRQC's deployment, the decisions made in 2025 could affect security in 2035 and beyond. Thirdly, the HNDL attack model implies that attackers do not have to wait with regard to the use of their own quantum capabilities to start enjoys harvesting encrypted OT communication today.

The Public Key Infrastructure (PKI) foundation for communications encrypted using TLS, such as remote access to SCADA systems, encrypted historian data and secure communication between distributed control nodes are fully dependent on RSA and ECC which are susceptible to Shor's algorithm. Wang et al. (2025) show that a quantum resistant network architecture will need to secure classical communication channels via PQC methods and will need orchestration across a diverse set of infrastructures, where it is important to carefully size key and signature requirements to work on constrained OT hardware.

3.3 Supply Chain and AI Model Integrity Risks

AI machinery driving critical infrastructure is also a third undervalued threat vector of the supply chain. AI models used in safety-critical systems can also be composite systems, built on top of foundation models, fine tuning data, inference libraries, orchestration libraries, and external APIs, all of which are potential attack surfaces. It also highlights that the field now lacks the tooling and practices that it has built around software Software Bill of Materials (SBOM) for traditional software components, as noted by Nocera et al (2025) by stating that comprehensive AI Bill of Materials (AIBOM) standards are not yet available. Operators are unable to systematically determine, if their deployments include which AI components, the provenance of these elements, and monitoring them to identify emerging vulnerabilities without AIBOM visibility. Critically, Mitera and Chowdhury 2024 suggest blockchain realizing SBOM and AIBOM sharing as a potential way to attain transparency in the supply chain, although vendor resistance of providing data on sensitive parts is a big practical issue. A summary of these attack vectors, affected sectors and current mitigation status is included in a consolidated threat matrix (Table 3).

Table 3: Critical Infrastructure Threat Matrix: Agentic AI and Quantum-Era Attack Vectors, Affected Sectors, and Mitigation Status

Threat Vector	Category	Affected Sectors	Attack Surface	Current Mitigation	Severity
Prompt Injection (Direct)	Agentic AI	All sectors	Agent input pipeline	Partial (input filters)	Critical
Prompt Injection (Indirect)	Agentic AI	Energy, Finance, Healthcare	External data sources	Minimal	Critical
NHI Credential Theft	Agentic AI	All sectors	Agent identity stores	Partial (PAM tools)	Critical
Multi-Agent Cascade Failure	Agentic AI	Energy, Water, Finance	Orchestration pipelines	Minimal	High
Tool-Call Hijacking	Agentic AI	Energy, Water, Transportation	OT API interfaces	Minimal	Critical
HNDL Attack	Quantum	All sectors	All encrypted comms	Minimal (pre-migration)	Critical
PKI/TLS Compromise	Quantum	All sectors	Internet-facing systems	Partial (hybrid PQC pilots)	Critical
OT Protocol Exposure	Quantum	Energy, Water, Transportation	DNP3, Modbus, IEC 61850	Minimal	High
AI Model Poisoning	Supply Chain	All sectors	Training pipelines	Minimal	High
AIBOM Gap Exploitation	Supply Chain	All sectors	AI component inventory	Minimal	Medium

Note. Severity ratings reflect assessed impact on critical infrastructure operations under successful exploitation. Current mitigation status reflects industry-wide posture as of Q1 2026. NHI = non-human identity; HNDL = harvest-now-decrypt-later; PKI = public key infrastructure; OT = operational technology; PAM = privileged access management; AIBOM = AI Bill of Materials.

4. Proposed Framework: AQIF

4.1 Design Principles

The Agentic-AI & Quantum-Resistant Infrastructure Framework (AQIF) is based on five overall operational design principles in keeping with the real-world experience of U.S. critical infrastructure and the dual threat domains that AQIF protects.

Least privilege and minimal footprint involves ensuring that each AI agent has only the absolute minimum permettant needed to accomplish its given task, limiting access to data, tools and services. Agent credentials should apply to limited set of resources and must have an expiration time and be revoked upon idle use. This principle directly tackles the NHI exploitation/lateral movement vectors identified in Section 3.

Continuous verification takes the Zero Trust principles to all agent-to-system and agent-to-agent interactions. All agents should be viewed as untrustworthy regardless of origin, prior authentication or network. Each tool use instance, access query for data and any communications between the agents should be authenticated and authorized with a policy engine continuously updated with behaviour indicators and signals, risk score from their context, and anomalies.

Any cryptographic implementation must now be designed as cryptographic agile meaning that any cryptographic system using it should be interoperable without redesign. This principle will be especially key when managing the PQC transition process, which will entail phase-outs of various heterogeneous infrastructure components and take a long period of time. Beginning with crypto-agility implementations is to support hybrid classical/post-quantum deployments as a step along the way, not as an end goal.

Human oversight and the ability to audit means that all consequent actions of agents must be logged, audit-able, and handled by a human. This includes actions that manipulate OT systems, give control commands or change security configurations. Automated systems for containment and roll back are needed to interrupt and roll back the actions of agents going beyond their scope of authority or showing abnormal behavior.

Resilience and graceful degradation assume that none of the security measures are invulnerable and that they can gracefully degrade in a possibly unsafe state if some of the security measures were compromised. Consider designing an agentic AI system to include boundaries to minimize blast radius, and consider a staged approach for migration to PQC to ensure failure of any single component does not negatively impact the cryptographic integrity of the system.

4.2 Layer 1 Agentic AI Workflow Security

The first layer of AQIF focuses on security measures from the moment of setting up the agent's identity to executing tasks, coordinating among agents, and handling incidents.

The Layer 1 control is non-human identity management. Each of them in an environment that involves a critical infrastructure should be assigned a unique verifiable identity that is controlled in the organization's identity governance system. A dedicated NHI management platform, separate from human identity management systems, should provision agent identities, automate credential rotation and continuously monitor for abnormal access patterns to an agent's credentials. Agent identities must be cryptographically bound to the software configuration of the agent, and by using attestation tools, discover unauthorized changes to the agent system's code and/or model weights.

The first line of defense against prompt injection attacks are prompt sanitization and input validation. Any external data that the agent ingests, such as what the user enters, documents retrieved from documents, API results, files from sensors, logging files, should go through a validation layer that does syntactic and semantic analysis of the data to identify adversarial instructions. Successful Defense Prompt Injection Rate is reduced from 73.2% to 8.7% while Baseline Task Performance is only 94.3% degraded in some of the defense methods documented by Li et al. (2025), which indicates that defense methods can be deployed without significant performance overhead. Architectural constraints should be used in addition to validation controls to introduce structural separation between instruction channels and data channels, and limit the "attack space" available to any indirect attempt at injection.

Constraining the scope of actions that an agent can perform allows restricting the set of tools and APIs that individual agents could invoke, so effectively enforcing the principle of least privilege at the operational level. Action scopes should be deployed at time of deployment based on the intended tool functional role and if an attempt is made to invoke an out of scope tool then alert and block invoked tool till human intervention. In an OT environment, one should clearly mandate that in general, such AIs are not permitted to give control orders without use of a preplanned, supervised and monitored human in the loop process.

Inter-agent communications are performed according to agents' authentication and authorization multi-agent trust protocols. Cryptographic authentication of agent-to-agent messages utilizing agent identity certificates, and an authorized agent registry at the orchestrator agents which refuses messages from unidentified peer agents and modified agents. To embed security, privacy and governance guarantees a priori to the design of multi-agent systems, Shah and Wang (2025) suggest taking inspiration from protocol engineering in distributed systems.

Continuous Runtime Visibility of Agent Actions with Behavioral monitoring & Anomaly detection. Behavioral baselines should be determined for each agent type, during the test and staging phases and include the following: Baseline data for each agent type on the following aspects: frequency of tool use, frequency of data access, volume of API calls, and inter-agent communication. These baselines should be used as benchmarks for production, and deviations monitored via AI-driven anomaly detection, feeding into escalated levels of warning. Lennon et al. (2025) underline that encompassing the entire agent lifecycle, from memory operations to retrieval-augmented generation (RAG) queries and orchestration of tool calls, is essential for detecting more advanced attacks that stay subtler than the sum of their constituent actions.

Incident response and agent containment procedures are the operational playbook used for incident responses to detected agent compromises. credential revocation, action blocking, agent quarantine should be automated actions taken in seconds after the detection of a compromise, and human escalation should be implemented for situations requiring judgment. Post incident forensic should hold a complete agent action log, including reasoning traces if they are available, which will aid the root cause analysis and improvement of the framework.

4.3 Layer 2 Quantum-Resistant Communication

AQIF layer 2 develops a formal migration plan for the transition from traditional cryptographic technologies to quantum-resistant communications infrastructure, particularly in OT/ICS settings.

The migration plan to NIST PQC is divided into three stages. During Phase 1 (0–18 months), organizations inventory their system, protocols and data flows relying on RSA or ECC, a cryptographic asset inventory. An order of priority is determined based on data sensitivity, system longevity and HNDL exposure. Beginning parallel deployment of hybrid classical/PQC configurations as

outlined below, in high priority systems, especially those with long-lived operational data and classified data. During Phase 2 (18-48 months), hybrid deployments cover most of the IT infrastructure, while OT-specific PQC implementations are trialled on individual or specific segments of IT infrastructure with sufficient computing capacity. Hardware Security Modules (HSMs) are purchased when cryptographic operations need to be hardware enforced (PQC-capable). As mentioned, the classical-only configurations of crypto will be decommissioned in Phase 3 (48–84 months), as all IT systems will be PQC deployed, and each OT system will have PQC migration if it has sufficient compute power. Legacy OT control systems, where space is a constraint to implement PQC, are cordoned off by quantum resistant encryption gateways.

Hybrid classical/PQC transitional deployment is used to acknowledge the fact that PQC migration can't be instantaneous. In a transitional period, systems could use a hybrid key establishment that would generally perform two key exchanges, one conventional (e.g., ECDH), and one PQC (e.g., ML-KEM per FIPS 203), and then use the two keys together to establish the session key. In this manner, the security of the session depends on the security of both the traditional and PQC schemes; an attacker may only breach the entire scheme if he breaches both of them. According to Rahman et al. (2025), the IETF has officially standardized hybrid PQC systems for important network protocols as SSH and TLS 1.3.

Certain strategies are needed to implement crypto-agility in an OT/ICS system, taking availability requirements, protocol rigidity, and resource constraints into account. Quantum-resistant encryption gateways can provide a protective wrapper, terminating external encrypted connections in PQC algorithms and then translating them to the legacies' native protocol inside the secure protection sphere of the OT network boundary. Finally, for new OT deployments and firmware changes, cryptographic agility must be an inherent design criterion and interfaces to the HWSMs must be compatible with support for changes in algorithms through firmware updates rather than HW changes. A detailed comparison of the three NIST-standardized PQC algorithms and their use in critical contexts of deploying infrastructure are presented in Table 4.

For the most critical communication links, including grid control networks, nuclear facility management networks, and classified government communication networks, which demand information-theoretic security, there is a strong desire to consider Quantum Key Distribution (QKD). As noted by Wang et al. (2025), deployments of QKD have shown exploitable vulnerabilities and given the situation where even computational security assumptions are not reliable, a layered security approach, with QKD over PQC, offer a defense in-depth.

Table 4: Comparison of NIST-Standardized Post-Quantum Cryptographic Algorithms (FIPS 203, 204, 205) and Their Applicability to Critical Infrastructure Communication

Algorithm	FIPS Standard	Cryptographic Basis	Public Key Size	Signature/ Cipher Size	Computational Overhead	OT/ICS Suitability	Primary Use Case
ML-KEM (Kyber)	FIPS 203	Module lattice (MLWE)	800–1568 bytes	768–1088 bytes (ciphertext)	Low–Medium	Moderate (resource-constrained devices need optimization)	Key encapsulation / TLS handshake
ML-DSA (Dilithium)	FIPS 204	Module lattice (MLWE/SIS)	1312–2592 bytes	2420–4595 bytes	Medium	Moderate (signature size may affect legacy protocols)	Digital signatures / code signing
SLH-DSA (SPHINCS+)	FIPS 205	Stateless hash-based	32–64 bytes	7856–49856 bytes	High (signing)	Limited (large signatures; best for infrequent signing)	Long-term document signing / firmware authentication

Note. Key and signature sizes reflect the range across security levels (2, 3, 5). Computational overhead is relative to RSA-2048 equivalent operations on contemporary hardware. OT/ICS suitability assessments assume deployment on modern embedded hardware; legacy PLCs and RTUs may require gateway-based implementations. ML-KEM = Module-Lattice Key Encapsulation Mechanism; ML-DSA = Module-Lattice Digital Signature Algorithm; SLH-DSA = Stateless Hash-Based Digital Signature Algorithm; MLWE = Module Learning With Errors; SIS = Short Integer Solution.

4.4 Layer 3 Integration of Governance and Compliance

Layer 3 consists of the governance framework and regulatory coordination to implement the technical controls through Layers 1 and 2 in the U.S. critical infrastructure policy sphere.

Regulatory Alignment provides an association between AQIF controls and the most important primary regulatory documents for each critical infrastructure sector. NERC Critical Infrastructure Protection (CIP) standards establish a baseline for the energy sector - NERC AQIF's NHI management and behavioral monitoring controls address requirements of NERC CIP-007 (Systems Security Management) and NERC CIP-010 (Configuration Change Management and Vulnerability Management). Compliance anchors in the financial sector are provided by Federal Financial Institutions Examination Council (FFIEC) guidance and Securities and Exchange Commission (SEC) cybersecurity disclosure rules. Technical safeguards required under HIPAA's Security Rule requirements are similar to those of AQIF, which includes access control and audit controls for healthcare. cISA's Secure by Design principles, a sector-agnostic guide released in November 2024, offer alignment points for all sectors, as do DHS AI safety recommendations for critical infrastructure.

SBOM requirements and AIBOM operationalize security in the supply chain in the governance layer. All AI systems deployed under the AQIF should have up-to-date AIBOMs that require them to document model provenance, training data sources, framework dependencies, and known vulnerabilities. AIBOM generation should be part of AI deployment pipelines, and AIBOMs should be remade whenever Model Weights, Framework Versions or External API Integrations change. Threat intelligence coordination is provided through AIBOM sharing with sector specific Information Sharing and Analysis Centers (ISACs).

In the context of red-teaming and adversarial testing protocols, agentic AI systems need to undergo regular testing under the various threat vectors outlined in Section 3. These exercises (red team exercises) in particular should be aimed at testing prompt injection resilience, exposure of NHI credentials, multi-agent cascades, and correctness of cryptographic implementations. Generative AI-powered Red-Teaming is rising to become best practice as reported by Ferrag et al. (2025), who highlight that AI-assisted Red-Teamers can explore Attack Surfaces systematically and at a scale and pace unattainable by human Red-Teamers.

The feedback mechanisms that keep AQIF effective as the threat landscape changes are continuity of monitoring and improvement. KPIs should include timeliness for detecting injections, anomaly score of agent action, cryptographic migration status, AIBOM coverage, and incident response times. Governance reviews may be performed on a prescribed schedule such as quarterly for operational KPIs and yearly for framework alignment, reflecting changes in regulatory landscape or significant changes in the threat landscape or NIST standards. The complete PQC migration roadmap is presented in a structured format in the form of a flowchart in Figure 3.

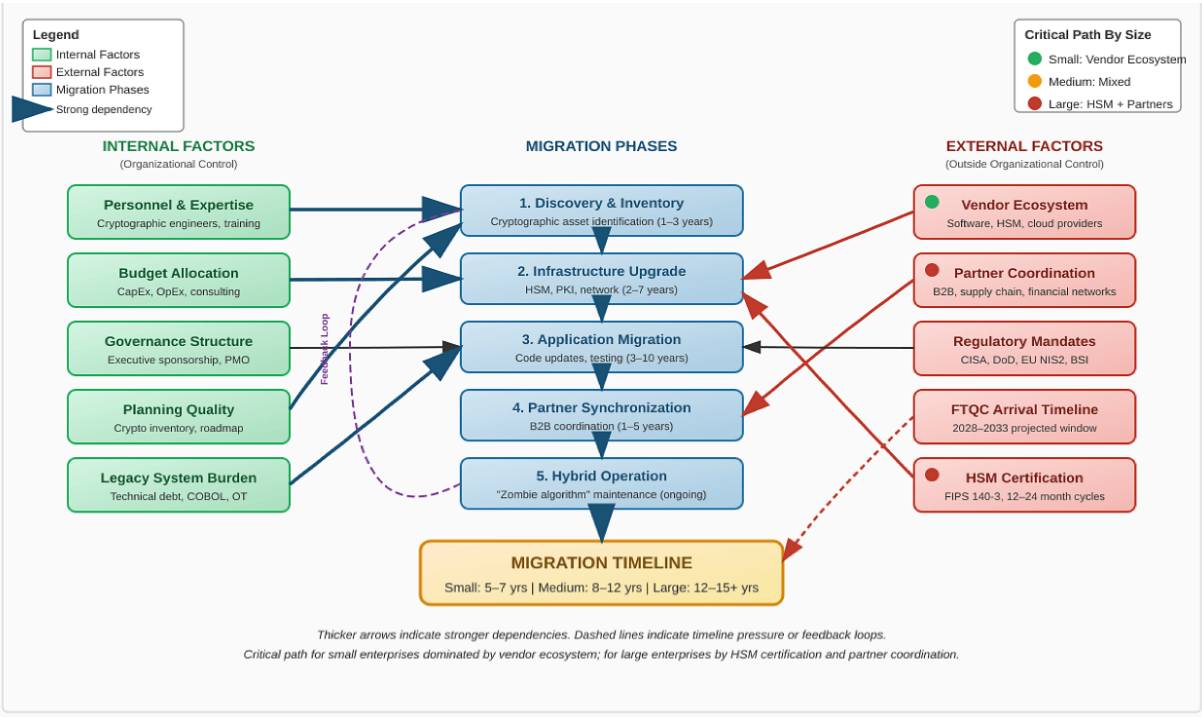

Figure 3. Migration Factor Dependency Diagram for Post-Quantum Cryptography Integration in U.S. Critical Infrastructure.

5. Discussion

5.1 Framework Effectiveness Against Identified Gaps

Based on the gaps highlighted from the literature study, the AQIF framework has incorporated both technical controls and governance structures which are not combined in other frameworks. Regarding security for agentic AI, AQIF's Layer 1 is the only operational control set specifically designed for deployment in critical infrastructure with agentic AI, which includes NHI management, prompt sanitization, multi-agent trust, and behavioral monitoring; capabilities which are not included in NIST CSF 2.0, the AI RMF, or Zero Trust SP 800-207. Specifically, with respect to quantum-resistant communication, AQIF's Layer 2 is a version of the NIST PQC standards that conveys the information to an operationally sequenced migration roadmap that considers the unique challenges of OT/ICS communications, which the NIST FIPS documents do not provide because they are focused on algorithm specification rather than operational deployment guidance.

The unique element of the framework is that it combines threat domains in a single governance structure. Before AQIF, an infrastructure operator attempting to implement guidance for addressing both AI security and PQC migration would have to schismatize guidance from at least six documents (NIST CSF 2.0, AI RMF 1.0, SP 800-207, FIPS 203, FIPS 204, FIPS 205) and manually address conflicts and gaps them arbitrarily. AQIF also ensures that there are clear cross-layer interdependencies for instance, the need to choose PQC-compatible digital signature algorithms from the beginning of issuing layer 1 identity credentials, rather than having them quantum vulnerable to begin with.

5.2 Comparison with Existing Frameworks

Twelve control dimensions have been systematically compared between AQIF and five existing frameworks in Table 5. The existing frameworks are shown to offer solid foundations in one or more dimensions, although they fail to cover the full range of twelve dimensions, including ZTA's continuous verification principles, CSF 2.0's governance structure, and AI RMF's lifecycle risk management. From a full coverage perspective, AQIF points to the full or partial coverage across all dimensions, its most important additions are the agentic specific threat controls (NHI management, prompt sanitization, multi-agent trust), OT specific PQC migration guidance and the integrated AIBOM requirements.

It should be noted that, AQIF is intended to be a supplementary framework and not a replacement to existing standards. Implementing AQIF should be along with existing regulatory requirements and infrastructures. The compliance crosswalk in Table 5 is provided to help operators understand how they encompass and/or go beyond existing requirements so less work is duplicated in implementing the AQIF controls.

5.3 Practical implications for operators and policy makers

AQIF is most valuable to infrastructure operators as a gap analysis. Table 5 demonstrates how other AQIF controls can be used by organisations that have already adopted NIST CSF 2.0 or Zero Trust controls to identify which controls address the missing agentic AI and PQC needs. The state of the world of quantum computing (QC) moves at a rapid pace, and there is a rapidly evolving migration roadmap for QC from NIST-recommended cryptographic inventories produced through NSM-10, to a sequence of operations.

In the policy arena, there are several areas where existing regulatory guidance might need to be updated, as noted below: NERC CIP standards were created before the advent of AI agents, hence the lack of coverage for AI behavioral monitoring and NHI management is growing more relevant as energy sector operators start to deploy AI agents in the grid management environment. In early 2026, CISA published guidance for OT environments, which involves a great step, but it is not yet agentic AI specific. The NIST Trustworthy AI in Critical Infrastructure Profile (under development) provides the best opportunity for embedding AQIF's Layer 1 controls under trusted guidance that can be adopted in all critical infrastructure sectors.

5.4 Limitations

AQIF has a number of important drawbacks. For one, it has not been tested through the large-scale rollout of it in production critical infrastructure environments. The threat matrix in Table 3, compliance crosswalk in Table 5 and migration roadmap in Figure 4 are derived from the current literature and expert judgment; not from operational data. The framework's specific recommendations may be further fine-tuned by empirical validation studies to compare the effectiveness and performance overhead of prompt sanitization controls for PQCs against new attack variants, in resource-constrained OT environments.

Secondly, the PQC migration roadmap in the framework takes for granted further advancements in the implementation of PQC algorithms for constrained devices. However, should current estimates prove incorrect with regard to the available resources, notably for legacy OT infrastructure that has very limited computing capability, the timeline for the Phase 3 may need to be stretched further and the quantum-resistant gateway approach may need to take more of a strain than originally planned.

Thirdly, there is an ever-changing threat landscape for agentic AI and quantum computing. As new attack techniques arise, AQIF's specific controls, especially those in Layer 1 will need to be subject of periodic review and adaptation. The governance processes in Layer 3 response with a continuous monitoring and framework review requirements provide for this evolution, which is only as effective as operators implementing these review cycles.

5.5 Future Research Directions

A few research pathways would materially enhance the evidence framework for AQIF and help drive progress of the overall field. Empirical studies regarding the deployment of agentic AIs on simulated critical infrastructure network settings could offer quantitative data on the success of the attacks, their time-to-detection and effectiveness in controlling them, which is lacking in the current theoretical analyses. Empirical comparative performance testing of NIST PQC algorithms on a variety of representative OT hardware platforms, including legacy PLCs, remote terminal units (RTUs), and field devices, would help to form more precise OT migration planning. Hammond et al.'s (2025) points out the lack of formal safety guarantees for agentic systems under adversarial conditions as a baseline open problem, which research in formal verification approaches for multi-agent security protocols would cover. Finally, sociotechnical studies on workforce readiness, organizational change management and implementation of governance for AQIF are needed to integrate the human and institutional aspects of deployment, which ultimately best predict successful implementation of the technical framework.

Table 5: Compliance Crosswalk: Proposed AQIF Framework Components Mapped to Existing U.S. Cybersecurity Standards and Regulations

AQIF Control	Layer	NIST CSF 2.0	NIST AI RMF	ZTA SP 800-207	NERC CIP	CISA Guidance	AQIF Coverage
NHI Identity Management	L1	Partial		Partial	Partial		Full
Prompt Sanitization	L1						Full
Constrained Action Scoping	L1	Partial	Partial	Full	Partial	Partial	Full
Multi-Agent Trust Protocols	L1			Partial			Full
Behavioral Anomaly Detection	L1	Partial	Partial	Partial	Partial	Partial	Full
Incident Response / Containment	L1	Full	Partial	Partial	Full	Full	Full
PQC Migration Roadmap	L2					Partial	Full
Hybrid Classical/PQC Deployment	L2					Partial	Full
OT/ICS Crypto-Agility	L2						Full
SBOM / AIBOM Requirements	L3	Partial				Partial	Full
Red-Teaming Protocols	L3	Partial	Partial		Partial	Partial	Full
Continuous Monitoring & Review	L3	Full	Full	Full	Full	Full	Full

Note. Full = the referenced standard provides comprehensive, actionable guidance aligned with the AQIF control. Partial = the standard provides related guidance that partially addresses the control. = the control is not addressed in the standard. L1 = Layer 1 (Agentic AI Workflow Security); L2 = Layer 2 (Quantum-Resistant Communication); L3 = Layer 3 (Governance and Compliance Integration). AQIF = Agentic-AI and Quantum-Resistant Infrastructure Framework; ZTA = Zero Trust Architecture; NERC CIP = North American Electric Reliability Corporation Critical Infrastructure Protection; CISA = Cybersecurity and Infrastructure Security Agency; NHI = non-human identity; AIBOM = AI Bill of Materials.

6. Conclusion

This combination of agentic AI adoption and quantum computers coming to maturity poses a twofold security challenge that matters deeply for U.S. critical infrastructure operators. New attack vectors such as prompt injection and multi-agent cascading failures and NHI exploitation are coming from agentic AI systems, which are offering transformative operational capabilities. At the same time, the harvest-now-decrypt-later threat model is rapidly becoming a reality, with adversaries archiving encrypted communications for decryption in the future.

In this paper Agentic-AI and Quantum-Resistant Infrastructure Framework (AQIF), a three-layer integrated security framework to tackle both threat domains simultaneously was introduced. Layer 1 ensures agentic AI workflow security via NHI identity management, prompting sanitization, limiting actions scope, multi-agent trust protocols, and behavioral monitoring. Building on Layer 2, it provides an OT aware and structured migration pathway to NIST-standardized post-quantum cryptographic algorithms with hybrid transitional deployment and crypto-agility design requirements. Layer 3 combines governance and compliance standards that fall in line with CISA, NERC CIP, the NIST AI RMF, and upcoming regulatory needs. The AQIF systematic compliance crosswalk analysis shows us that based on implementation alone, they individually don't address the gaps.

It is integration that forms the main value of the framework. However, the technical controls for success in agentic deployment and communications infrastructure quantumization are not orthogonal; agent identity credentials need to be quantum resistant, PQC migration governance should provide quantum-resistant threat detection for AI systems, and supply chain security should double down on both AI model security and cryptographic component security. The unified architecture is provided by AQIF in which these interdependencies are explicitly managed.

The article's findings suggest a number of practical steps for policymakers: urgently update NERC CIP standards to incorporate agentic AI present in energy OT environments; make it a mandatory requirement to comply with AIBOM requirements for AI systems deployed in federally regulated critical assets; and accelerate the development of the NIST Trustworthy AI in Critical Infrastructure Profile for authoritative Sector-Agnostic Guidance for security controls to be used with agentic AI systems. Given the nature and the challenges involved in operating in a legacy OT environment, and the governance nature of large infrastructure environments, the phased implementation roadmap offers a structured entry point for infrastructure operators.

The framework needs to be validated through its practical application and implementation, and as new uses of AI and agent technology continue to emerge, and novel quantum threats become evident, the framework must be continuously reassessed and adapted. None of these are detriments that reduce the value of the framework; each is a quality of a security framework on the cutting edge of technological transformation. What AQIF offers is a structured integrated base upon which this work can build.

Funding: This research received no external funding.
Conflicts of Interest: The authors declare no conflict of interest.
Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers.

References

[1]. Agrawal, A., & Mishra, R. (2025). Transforming cybersecurity with agentic AI to combat emerging cyber threats. Computers & Security, 143, 103985. https://doi.org/10.1016/j.cose.2025.103985

[2]. Al Kuwaiti, M., & Ismail, H. (2026). Adversarial attacks on large language models: A survey. In V. Bhateja, M. El Barachi, A. T. Azar, & D. K. Sharma (Eds.), Information system design: Big data analytics and data science. ISDIA 2025 (pp. 401–415). Springer. https://doi.org/10.1007/978-981-96-9248-4_40

[3]. Alladi, T., Chamola, V., Parizi, R. M., & Choo, K.-K. R. (2025). Zero-trust industrial network security using AI and explainable inference. World Journal of Advanced Research and Reviews, 28(2), 1136–1154. https://doi.org/10.30574/wjarr.2025.28.2.3705

[4]. Chen, X., & Liu, Y. (2025). Post-quantum cryptography and quantum-safe security: A comprehensive survey. arXiv. https://arxiv.org/abs/2510.10436
[5]. Ferrag, M. A., Tihanyi, N., Cordeiro, L. C., & colleagues. (2025). Generative AI in cybersecurity: A comprehensive review of LLM applications and vulnerabilities. Future Generation Computer Systems, 173, 107877. https://doi.org/10.1016/j.future.2025.107877
[6]. Hammond, L., Chan, A., Clifton, J., Hoelscher-Obermaier, J., Khan, A., McLean, E., Smith, C., Barfuss, W., Foerster, J., Gavenciak, T., & Han, T. A. (2025). Multi-agent risks from advanced AI (Technical Report No. 1). Cooperative AI Foundation. https://doi.org/10.48550/arXiv.2502.14143
[7]. Hassan, M., & colleagues. (2025). Quantum-secure data centers: Preparing critical infrastructure for the post-quantum era. World Journal of Advanced Research and Reviews, 26(2), 2950–2959. https://doi.org/10.30574/wjarr.2025.26.2.1742
[8]. Jain, S., & Singh, P. (2025). A systematic literature review on the implementation and challenges of zero trust architecture across domains. Electronics, 14(5), 1039. https://doi.org/10.3390/electronics14051039
[9]. Kaur, H., Sharma, V., & Patel, N. (2025). Generative AI and LLMs for critical infrastructure protection: Evaluation benchmarks, agentic AI, challenges, and opportunities. Sensors, 25(6), 1666. https://doi.org/10.3390/s25061666
[10]. Lee, D., & colleagues. (2024). Prompt infection: LLM-to-LLM prompt injection within multi-agent systems. arXiv. https://arxiv.org/abs/2410.07283
[11]. Lennon, D., Datta, A., & colleagues. (2025). TRiSM for agentic AI: A review of trust, risk, and security management in LLM-based agentic multi-agent systems. AI Open, 6, 100212. https://doi.org/10.1016/j.aiopen.2026.100212
[12]. Li, Y., Zhang, X., & Chen, W. (2025). A survey of agentic AI and cybersecurity: Challenges, opportunities and use-case prototypes. arXiv. https://arxiv.org/abs/2601.05293
[13]. Liu, Z., Chen, H., & Wang, F. (2025). Large language models in cybersecurity: A survey of applications, vulnerabilities, and defense techniques. AI, 6(9), 216. https://doi.org/10.3390/ai6090216
[14]. McHugh, J., Šekrst, K., & Cefalu, J. (2025). Prompt injection 2.0: Hybrid AI threats. arXiv. https://arxiv.org/abs/2507.13169
[15]. Mitra, S., & Chowdhury, R. (2024). Trust in software supply chains: Blockchain-enabled SBOM and the AIBOM future. In Proceedings of the 2024 ACM/IEEE 4th International Workshop on Engineering and Cybersecurity of Critical Systems (pp. 1–10). ACM. https://doi.org/10.1145/3643662.3643957
[16]. Nocera, S., Penta, M., Ahmed, F., Romano, S., & Scanniello, G. (2025). What we know about AIBOMs: Results from a multivocal literature review on artificial intelligence bill of materials. ACM Transactions on Software Engineering and Methodology. Advance online publication. https://doi.org/10.1145/3786773
[17]. Paulraj, J., Raghuraman, B., Gopalakrishnan, N., & Otoum, Y. (2025). Autonomous AI-based cybersecurity framework for critical infrastructure: Real-time threat mitigation. arXiv. https://arxiv.org/abs/2507.07416
[18]. Rahman, A., Hossain, M., & Islam, T. (2025). Post-quantum cryptography in practice: A literature review. IACR ePrint Archive. https://eprint.iacr.org/2025/1668
[19]. Rashed, M., & colleagues. (2026). AI-enabled cybersecurity framework for future 5G wireless infrastructures. Scientific Reports, 16, 7821. https://doi.org/10.1038/s41598-026-37444-8
[20]. Rodriguez, A., & Martinez, C. (2025). Open challenges in multi-agent security: Towards secure systems of interacting AI agents. arXiv. https://arxiv.org/abs/2505.02077
[21]. Sarker, I. H. (2024). AI-driven cybersecurity and threat intelligence: Cyber automation, intelligent decision-making and explainability. Springer Nature. https://doi.org/10.1007/978-3-031-54534-4
[22]. Scrivano, A. (2025). A comparative study of classical and post-quantum cryptographic algorithms in the era of quantum computing. arXiv. https://arxiv.org/abs/2508.00832
[23]. Shah, R., & Wang, L. (2025). Security considerations for multi-agent systems. arXiv. https://arxiv.org/abs/2603.09002
[24]. Singh, Y., Patel, N. D., & Shandilya, S. K. (2024). Large language models for cyberattack defense: A critical survey. Knowledge and Information Systems. Advance online publication. https://doi.org/10.1007/s10115-026-02736-y
[25]. Tanaka, K., & Okonkwo, C. (2025). Prompt injection attacks in large language models and AI agent systems: A comprehensive review of vulnerabilities, attack vectors, and defense mechanisms. Information, 17(1), 54. https://doi.org/10.3390/info17010054
[26]. U.S. Department of Commerce, National Institute of Standards and Technology. (2024a). FIPS 203: Module-lattice-based key-encapsulation mechanism standard. https://doi.org/10.6028/NIST.FIPS.203
[27]. U.S. Department of Commerce, National Institute of Standards and Technology. (2024b). FIPS 204: Module-lattice-based digital signature standard. https://doi.org/10.6028/NIST.FIPS.204
[28]. U.S. Department of Commerce, National Institute of Standards and Technology. (2024c). FIPS 205: Stateless hash-based digital signature standard. https://doi.org/10.6028/NIST.FIPS.205
[29]. Wang, J., & colleagues. (2025). Quantum-resistant networks using post-quantum cryptography. arXiv. https://arxiv.org/abs/2510.24534
[30]. Xu, H., Liu, Y., Xing, Y., & colleagues. (2025). Large language models for cyber security: A systematic literature review. ACM Transactions on Software Engineering and Methodology, 34(2), 1–39. https://doi.org/10.1145/3698399