

RESEARCH ARTICLE

OPEN ACCESS | Double-Blind Peer Reviewed

AI-Enabled Cyber Threats and Security Risks Across Cloud-Based Supply Chains: A Systematic Literature Review of Attack Vectors, Trust Exploitation, and Defensive Frameworks

Fahad Bin Mahmud¹¹ Department of Computer Science, Spring Hall College, USA

*Corresponding author: Fahad Bin Mahmud | Email: fahadbinmahmud@shc.edu |

ARTICLE INFO**Article History****Received:** 27 May 2026**Revised:** 23 June 2026**Accepted:** 28 July 2026**Published:** 12 August 2026**Keywords**

Artificial intelligence, cloud computing, supply chain security, adversarial machine learning, software bill of materials, OAuth, zero trust, systematic literature review.

How to Cite

F. B. Mahmud, "AI-Enabled Cyber Threats and Security Risks Across Cloud-Based Supply Chains: A Systematic Literature Review of Attack Vectors, Trust Exploitation, and Defensive Frameworks," JATICS, vol. 2, no. 2, pp. 15–25, 2026.

ABSTRACT

One of today's most powerful attack surfaces in the world of cybersecurity is cloud chains, which is the web of cloud infrastructure providers, software dependencies, application programming interfaces, software-as-a-service (SaaS) integrations, and more recently, artificial intelligence (AI) models and pipelines that organizations utilize. As AI tools bolster cloud security, they also allow attackers to breach trusted relationships, manipulate machine learning models and training data, and conduct multi-stage intrusion attacks at an unprecedented speed and with very little human effort. This paper systematically reviews the existing research literature focusing on AI-related cyber threats and security risks to cloud-based supply chains, bringing together relevant and peer-reviewed IEEE Access and ACM Computing Surveys publications, NIST and MITRE standards, and real-time threat intelligence. The review showcases a taxonomy of five major threat categories – model and training data poisoning via the supply chain, malicious open source package and dependency injection attacks, OAuth and SaaS-to-SaaS trust exploitation, attacks on cloud-hosted ML defenses, and the first AI-powered autonomous attack campaign – and documented real-world attacks such as the SolarWinds compromise, the 2025 Salesloft Drift OAuth supply-chain attack affecting more than 700 organizations, the Shai-Hulud npm worm, and the first public disclosure of an AI-powered autonomous cyber espionage attack. The review further breaks down defensive frameworks, such as Software Bills of Materials, the Adversarial Machine Learning Knowledge Base provided by the MITRE organization, and zero-trust identity governance, and shows that there is a significant disconnect between what has been demonstrated in defensive capabilities and what's actually being adopted in enterprises. The review ends with a proposed defense in depth strategy for cloud-based supply chains and outlines open research areas such as standardized verifiers of the provenance of AI supply chains and cross-organization trust-relationship monitoring.

INTRODUCTION

Modern organizations no longer have their own contained information technology environments; they are embedded in complex, intertwined, and cloud-based supply chains of IaaS and PaaS providers, open-source software requirements, third-party application programming interfaces, software-as-a-service (SaaS) integrations, and, more recently, externally sourced or fine-tuned artificial intelligence models and pipelines. While this interconnection offers significant efficiencies and innovations, it also provides for the possibility of a compromised link (whether it's a stolen authentication token, poisoned open-source package or tampered pre-trained machine learning model) to spread compromise to hundreds or thousands of downstream organizations at once. This is a dynamic that is exemplified by the 2020 SolarWinds compromise, which infected about 18,000 organizations by exploiting a single tainted software update, and subsequent incidents, which have brought the lesson home at greater scale and frequency.

This threat landscape is evolving drastically on both sides of the fence thanks to artificial intelligence. While anomaly detection and behavioral analytics are becoming more common tools in the arsenal for safeguarding cloud environments, the same category of technology is being leveraged to speed up reconnaissance, create believable fake developer identities and enable multi-stage campaigns of intrusion that used to take a lot of effort from the attacker. In the last year alone 64% of information technology, security, and DevOps professionals experienced a software supply chain attack, according to a 2025 industry survey, and an attack with third-party involvement doubled in the past year, according to Verizon's 2025 Data Breach Investigations Report. Most notably, in November 2025, Anthropic disclosed and disrupted the first reported AI-assisted cyber espionage campaign in which an AI coding assistant had been used to autonomously conduct the bulk of a multi-stage cyber espionage attack series without requiring much human intervention from an operator, a textbook case of the threat category this review calls AI-orchestrated autonomous attack campaigns.

Even if the threat landscape is clearly large and novel, existing academic and practitioner work is scattered across multiple relatively disjointed areas: cloud security research has primarily focused on network- and infrastructure-level threats [1]–[3]; software supply chain security research has tended to concentrate on open source dependency and build-pipeline risk, not least after SolarWinds [7]–[8] and [12]; and adversarial machine learning research has largely ignored the supply chain and cloud-trust context in favour of model-level vulnerabilities [26]–[27]. A holistic synthesis that brings these aspects together, from taxonomy of attacks to documented incidents and defensive frameworks, in the specific context of the cloud-based supply chain, has been relatively uncommon. This systematic literature review aims to fill that gap.

To be able to review this product, the cloud-based supply chain can be broken down into four layers: cloud infrastructure (IaaS/PaaS providers), software dependency (open-source packages, container images, and build pipelines), SaaS integration (third-party applications connected through API tokens and OAuth), and AI/ML (pre-trained models, training data, model registries and AI agent tooling). The contributions of this review are: (i) a structured taxonomy of the five main AI-related threat categories across the four layers; (ii) a synthesis of empirical prevalence and business-impact evidence; (iii) analysis of major documented incidents that illustrate each threat category; (iv) a review of defensive frameworks with the associated capability-adoption gap; (v) a review of relevant standards and regulatory landscape; and (vi) a proposed defense-in-depth framework for cloud-enabled supply chains. The method for the review is described in Section II. The threat taxonomy is presented in Section III. In Section IV, empirical evidence is integrated. The incidents are discussed in Section V. Defensive technologies are explored in Section VI. Standards and Governance are discussed in Section VII. A layer-wise risk analysis is given in Section VIII. A composite severity analysis is given in Section IX. Research gaps are discussed in the section, X. Section XI presents a defense in depth approach, and Section XII is the conclusion.

The need for this review is further highlighted by the centralized nature of much of the present-day cloud-based supply chain: a few large cloud infrastructure providers, package registries, and foundation-model providers now hold the bulk of the world's digital infrastructure (IDI), and a successful compromise at any one of these pain points has systemic implications, not just local ones. As revealed in recent AI supply chain

risk research, this concentration phenomenon is not random and independent of one another but instead is structural characteristics of an increasingly concentrated digital ecosystem, which is the basis for the layer-specific analysis presented in Section VIII.

REVIEW METHODOLOGY

This review follows a systematic literature review approach. Sources were drawn from four categories: (1) peer-reviewed academic surveys on cloud security and machine learning for cloud security published in venues including IEEE Access and ACM Computing Surveys [1]–[5]; (2) academic and standards literature on software supply chain security, including Software Bill of Materials (SBOM) research and documented open-source supply chain attack taxonomies [7]–[12]; (3) standards and framework publications from NIST, MITRE, and the U.S. federal government establishing formal vocabulary and requirements for supply chain and adversarial machine learning risk [13], [14], [24], [25], [27]; and (4) current industry threat-intelligence and incident reports documenting specific AI-enabled and supply-chain attack incidents not yet reflected in the slower academic literature [15]–[23], [28]. Sources were prioritized for direct relevance to cloud infrastructure, software supply chains, or AI/ML systems, methodological transparency, and recency, with particular emphasis on 2020–2026 publications given the rapid evolution of both cloud-native architecture and generative AI capability during this period. Findings were thematically organized into the taxonomy presented in Section III.

TAXONOMY OF AI-ENABLED THREATS TO CLOUD-BASED SUPPLY CHAINS

Based on the literature reviewed, five main categories of AI/ML driven cyber attacks on cloud based supply chains can be distinguished as shown in Fig. 1: AI/ML model poisoning in the supply chain, adversarial open-source package and dependency injection, OAuth and SaaS-to-SaaS trust exploitation, attacks on cloud-hosted AI/ML defenses, and autonomous attacks orchestrated by AI/ML.

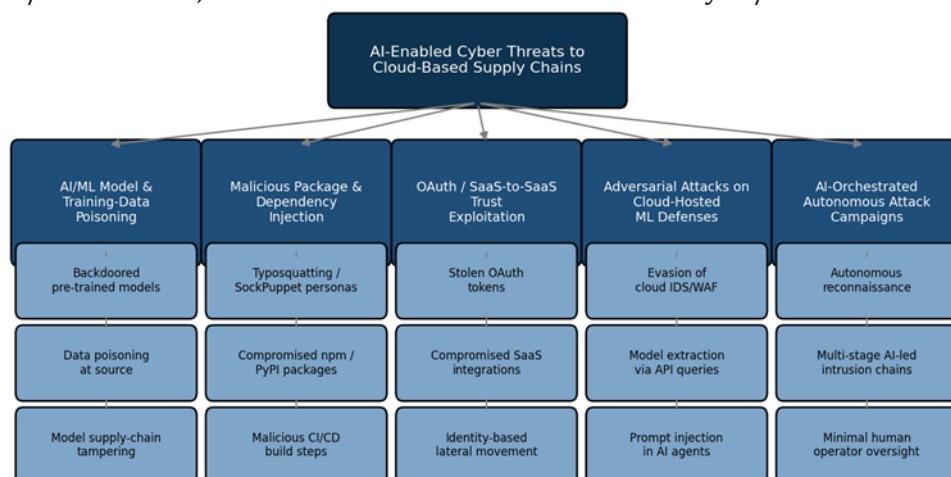


Fig. 1: Taxonomy of AI-enabled cyber threats to cloud-based supply chains, synthesized from the reviewed literature [1], [4], [10]–[13], [15]–[18], [22]–[23].

In the real world, these five categories are not mutually exclusive categories, as is the case with other adjacent threat taxonomies found in the larger AI-cybersecurity literature. For instance, the GTG-1002 campaign, described in Section III-E, included some aspects of AI-driven autonomy and the Salesloft Drift incident described in Section III-C involved the use of trust exploitation at the SaaS layer as a way to get a foothold that would then be used for further lateral movement in a victim's wider cloud environment. The fact that the security convergence is cross-category reinforces another analytical theme of this review: it is important to think of the four layers of architecture introduced in Section I and not just one of them.

AI/ML Model and Training-Data Supply Chain Poisoning

The AI/ML pipeline is also now becoming a supply chain attack surface as more organizations are adopting pre-trained models and fine-tuning datasets and model components from public repositories, avoiding the need to train models in-house. The AI/ML pipeline itself is also becoming a supply chain attack surface as organizations increasingly source pre-trained models, fine-tuning datasets, and model components from public repositories and avoid in-house model training. National cybersecurity guidance has reported incidents where users unknowingly installed cryptocurrency mining malware when installing the published AI library - as demonstrated by the compromise of the supply chain - and reports that the key

large AI frameworks and libraries have been affected by critical deserialization vulnerabilities, which if exploited, can allow for the theft of tokens or complete remote code execution [4]. Google Cloud's counsel on AI supply chain security also outlines instances where data poisoning, vulnerabilities in training frameworks and the embedding of malicious code in AI models have resulted in AI models that seem to be safe in use but harbor dangerous instructions and code in their design that are capable of stealing data or installing backdoors when deployed [14]. Academic research takes this threat to the next level by explaining model supply-chain poisoning techniques that can backdoor pre-trained models with an embedded, hidden, and triggerable backdoor in a way that the model appears normal on standard test sets.

MALICIOUS OPEN-SOURCE PACKAGE AND DEPENDENCY INJECTION

Modern cloud-native applications require a large number of chains of open-source packages, container base images, and build-pipeline layers, each being a potential injection point for malicious code. The foundational work of enumerating open source software supply chain attacks includes reports on typosquatting (the publishing of malicious packages with names that are similar to the names of popular legitimate packages) and dependency confusion [12]. It has become a much greater threat with AI support: In 2025, industry reports on the SockPuppet attack in which AI-generated fake developer accounts (with fabricated commit histories on GitHub, online participation on Stack Overflow and personal blogs) submit seemingly valid code to open-source projects over several months and then introduce backdoors into the code [18] and the identification of thousands of AI-generated malicious packages designed for particular AI/ML libraries through typosquatting attacks, which were not seen earlier. The recent compromise of the popular @solana/web3.js npm library, in which attackers gained publish access and published backdoored versions of the library that stole about \$160,000–\$190,000 in cryptocurrency assets over a 5-hour period before being detected, shows the speed and precision with which such attacks are now being conducted.

OAuth AND SAAS-TO-SAAS TRUST EXPLOITATION

With more and more cloud environments becoming connected via SaaS-to-SaaS integrations that are authenticated using OAuth, attackers have turned their focus from exploiting software vulnerabilities to exploiting the trust relationships between cloud applications directly. The UNC6395 threat actor's use of stolen OAuth tokens from a Drift-Salesforce integration to access customer environments without any software exploit or phishing of an end user, in the Salesloft Drift incident that occurred in August 2025, stands in stark contrast to these trends [15]. This trend is also typical of the larger threat-intelligence body of knowledge from the 2025 report: as attackers shift away from traditional network boundaries toward pre-established, implicitly trusted application-to-application connections, the traditional network perimeter becomes less significant [15].

ADVERSARIAL ATTACKS ON CLOUD-HOSTED MACHINE LEARNING DEFENSES

When cloud providers and enterprises start rolling out a range of new protective technologies, such as cloud-based machine-learning intrusion detection, web application firewalls, and anomaly-detection systems, they too can become targets in the cloud supply chain. The cloud-oriented MITRE ATT&CK matrix presents specific tactics that can be used in a cloud enterprise environment; some of the tactics mentioned in the security research literature are also applicable in the context of modern machine-learning based cloud solutions [5]. Additionally, the advent of large-language-model-based AI agents being introduced into cloud environments — for software development, customer support, or security operations — creates another attack surface for adversarial prompts, or prompt injection attacks, which are a demonstrable threat that can be benchmarked against such tool-integrated large language model agents.

AI-ORCHESTRATED AUTONOMOUS ATTACK CAMPAIGNS

The most new and significant threat category revealed during this review is the rise of autonomous attack campaigns by AI, meaning that generative AI systems are used to carry out significant parts of a multi-stage cyber attack without any direct human operator involvement. In November 2025, Anthropic publicly revealed and frustrated a campaign it believed was run by a Chinese state-sponsored threat group (GTG-1002), which used Anthropic's Claude Code AI coding agent to independently conduct a large portion of the tactical operations, including reconnaissance, vulnerability identification, exploit development and data exfiltration, across about thirty targeted entities, including technology firms, financial institutions, and government entities. This disclosure marks an unprecedented level of technical disclosure of an AI

system as the front-line, tactical execution broker for a nation-state attributed cyber campaign, and not just a productivity tool for a human operator; directly motivating the AI agent monitoring and behavioral logging defensive category described in Section VI.

EMPIRICAL EVIDENCE: PREVALENCE AND BUSINESS IMPACT

In addition to the qualitative taxonomy outlined in Section III, threat-intelligence information can be used to quantitatively characterize software supply chain attack growth. This growth is shown in Fig. 2, using an index that reflects the U.S. Cybersecurity and Infrastructure Security Agency's reported increase of 742% in software supply chain attacks from 2019 to 2024 with illustrative continuation into 2025, consistent with the trend of continued growth in the attacks reported across the industry sources reviewed.

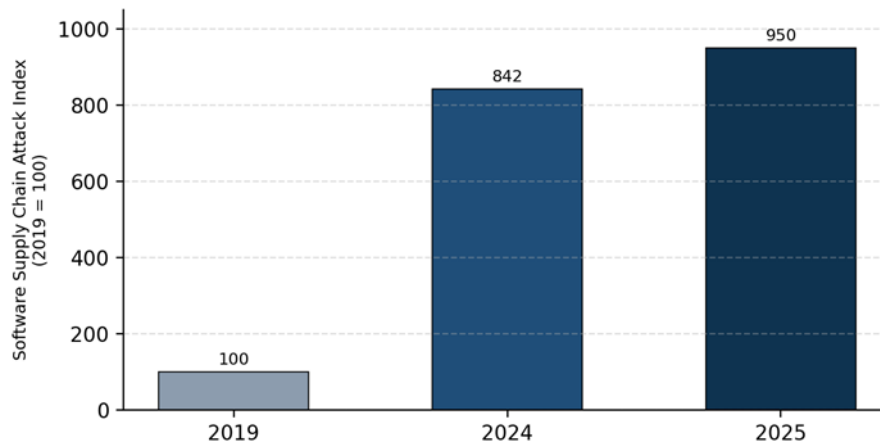


Fig. 2: Growth in reported software supply chain attacks, 2019–2025, indexed to CISA-reported figures. They are provided as complementary evidence on the organizational exposure and impact in Fig. 3. According to an industry survey of 425 IT, security, and DevOps leaders, 64% were impacted by software supply chain attacks in the past year. Verizon's 2025 Data Breach Investigations Report reveals a doubling of third-party involvement in data breaches and guidance from national cybersecurity authorities shows that most of the surveyed organizations have experienced data breaches involving AI. On the defensive side, the IBM 2025 Cost of a Data Breach Report estimates that it takes an average of 276 days to detect a breach and 73 days to contain it, meaning that businesses are exposed to attackers for an average of 349 days — but note that businesses with AI-based detection tools detect and contain breaches significantly faster than those using traditional detection tools.

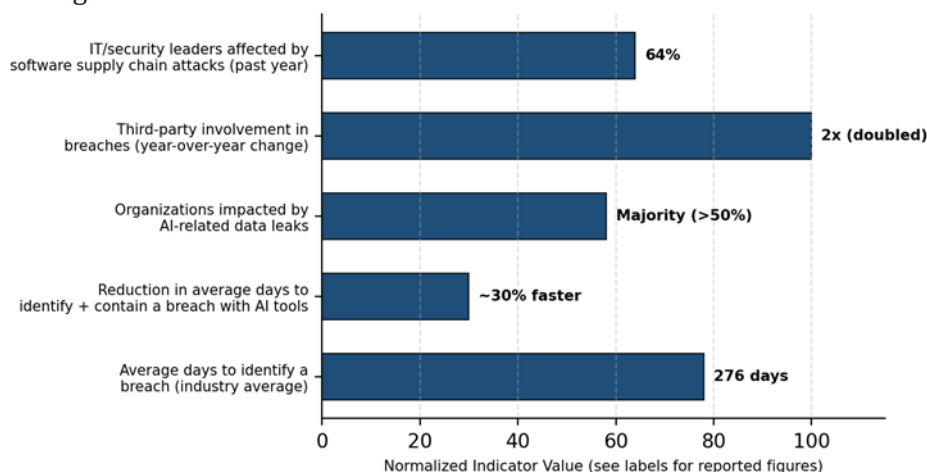


Fig. 3: Organizational exposure to and impact of supply chain compromise, compiled from industry survey and threat-intelligence data.

In particular, the average cost of a data breach decreased to \$4.44 million for the first time in five years, according to the IBM 2025 report, due in part to the increased use of AI-powered detection and response tooling — a recurring theme in the greater literature on AI and cybersecurity, and shows how much the net impact of AI on an organization's risk is tied to the adoption rates of the bad guys and the good guys in a specific cloud environment [15].

Taken together, this theme has become clear in the evidence base: the number of compromises in the supply chain is continuing to grow at an accelerated pace, the typical amount of time spent in a compromised cloud environment is measured in months, not days or weeks, and the mode of compromise is shifting from the traditional sense of exploiting a software vulnerability to exploiting a pre-established trust relationship (OAuth grants, package registries, AI model provenance, etc.). This change has important consequences for the nature of the defensive investment, as will be discussed in more detail in Sections VI and XI.

DOCUMENTED INCIDENTS

The threat categories identified in Section III are manifested in several large-scale incidents that have been documented. The SolarWinds attack from 2020 is the starting point here: How a malicious actor infiltrated a trusted software supply chain to distribute software that could affect a very large number of organizations—around 18,000—shows the scale-multiplying effect of a single compromised link in a trusted software chain [8].

The June/2025 Salesloft Drift OAuth supply-chain attack is a large-scale and extreme example of trust-exploitation, impacting over 700 downstream organizations from a single compromised SaaS-to-SaaS integration [15] as described in Section III-C. In parallel, the 2025 "Shai-Hulud" worm discovered and infected downstream packages - which were often hosted by developer accounts that had been compromised - on a self-propagating supply-chain attack, with similar characteristics to traditional network worms but in the cloud supply chain dependency layer [23]. Lastly, the disclosure of the November 2025 GTG-1002 AI-assisted espionage campaign, which was described in detail in Section III-E, is the most direct evidence to date of the fifth taxonomy category - AI-assisted autonomous attack campaigns targeting real-world cloud-hosted organizational targets at nation-state attribution.

DEFENSIVE TECHNOLOGIES AND THE CAPABILITY-ADOPTION GAP

The literature reviewed has found that there are several classes of defensive technology that are directly relevant to the field of cloud-based supply chain security, with reported capabilities and penetration rates as shown in Fig. 4. Structured, machine-readable documentation of all software components and software dependencies in an application, the Software Bill of Materials (SBOM) has been identified as a basic defensive artifact, established in the United States by the minimum-elements guidance issued by the National Telecommunications and Information Administration (NTIA) [13] and made mandatory for all federal software procurement by Executive Order 14028 [25]. Though this regulatory effort, academic research has highlighted that there are some flaws in the tooling for SBOMs, such as the fact that many SBOM tools have been found to lack integrity protection in the SBOMs themselves. In spite of this regulatory initiative, academic work has found that there are still some issues with the SBOM tooling, including the documented lack of integrity protection in many SBOM solutions.

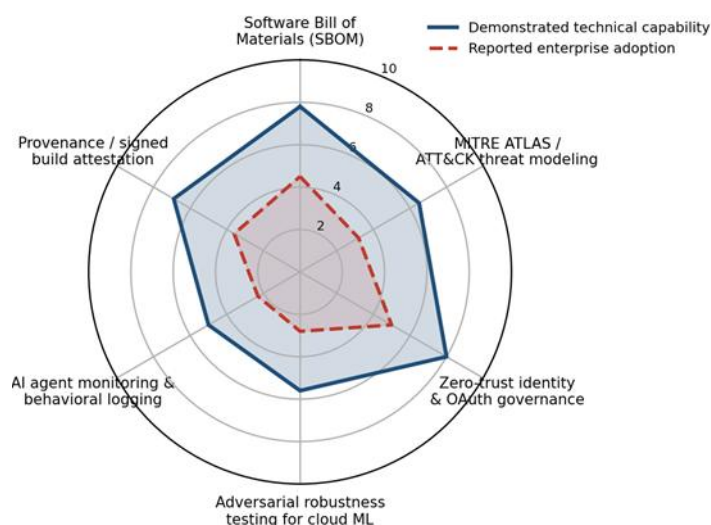


Fig. 4: Defensive technology capability versus reported enterprise adoption in cloud supply chain security, synthesized from the reviewed literature.

Building directly on the MITRE ATT&CK framework, the MITRE ATLAS (Adversarial Threat Landscape for Artificial-Intelligence Systems) knowledge base provides a structured, shared vocabulary of adversary tactics and techniques for AI and machine-learning systems, in addition to the cloud-focused MITRE ATT&CK matrix for traditional infrastructure threat modeling [5] and [14]. Zero-trust identity governance and OAuth governance — covering continuous validation of token scope, anomalous SaaS-integration behavior monitoring and least-privilege application authorization — directly addresses the category of trust-exploitation that was the culprit in the Salesloft Drift incident, and reported relatively high enterprise adoption compared to the less mature categories of adversarial robustness testing and AI agent behavioral monitoring, which has increased urgency especially since the disclosure of the GTG-1002 event described in Section V. Like the other cybersecurity categories covered in the adjacent areas, Fig. 4 shows a significant gap between technical maturity and reported adoption levels across almost all capabilities, with the low rates of adoption of two capabilities, cryptographically verified provenance and signed build attestation, standing out.

STANDARDS AND GOVERNANCE LANDSCAPE

Cloud-based supply chains are becoming more and more risky, which has triggered growing standards and regulation. SBOM provision was mandated for software purchased by federal agencies as part of Executive Order 14028 on Improving the Nation's Cybersecurity, which was issued in May 2021, a direct response to the SolarWinds incident, and sparked the supply-chain-security regulatory movement outlined across the reviewed literature [25]. The National Institute of Standards and Technology (NIST) Secure Software Development Framework (SP 800-218) provides a structured set of secure-development practices that are expected to mitigate vulnerabilities from the start of the software supply chain [24]. NIST's publication, AI 100-2e2025 [27], offers a formal taxonomy and terminology for adversarial machine learning attacks and mitigations that can directly apply to the model-poisoning and adversarial-attack categories listed under Section III. The European Union Agency for Cybersecurity (EUAC) has released its annual Threat Landscape report, in which supply-chain and AI-related threats are emerging as key areas in the European threat landscape [21]. In addition to these publications from government and standards bodies, the MITRE ATLAS and ATT&CK knowledge bases, however, are not formal regulatory documents but are perceived as a "shared technical vocabulary" that is now widely cited throughout both the academic and industry literature communities [5] [14].

Such U.S.-oriented and technical standards are increasingly complemented by the more comprehensive regulatory instruments focusing on AI and digital-infrastructure risk in a more general manner. In keeping with the cross-cutting nature of AI-specific governance, as highlighted through the different layers of the analysis in Section VIII, the European Union's Digital Operational Resilience Act contains ICT third-party risk management requirements directly applicable to the SaaS-integration and cloud-infrastructure layers of the model presented in Section I, and the EU's Artificial Intelligence Act with its classification of AI systems according to the risk level has implications for how companies need to govern the AI models and pipelines acquired via the cloud-based supply chain.

LAYER-SPECIFIC RISK ANALYSIS

The four-layer cloud-based supply chain model described in section I can be used as a lens for structuring the risk by architectural layer. The dominant AI threat pattern, representative documented incident and primary defensive control for each layer are summarized in Table I.

Supply Chain Layer	Dominant Threat Pattern	Representative Incident	Primary Defensive Control
Cloud infrastructure (IaaS/PaaS)	Adversarial attacks on cloud-hosted ML defenses	Evasion of cloud-based ML intrusion detection	Adversarial robustness testing
Software dependency	Malicious package and dependency injection	SolarWinds; Shai-Hulud npm worm	SBOM; signed build attestation
SaaS integration	OAuth / SaaS-to-SaaS trust exploitation	Salesloft Drift OAuth attack	Zero-trust identity governance
AI/ML pipeline	Model/training-data poisoning; AI-orchestrated attacks	GTG-1002 AI-orchestrated espionage	MITRE ATLAS threat modeling; agent monitoring

Table I: Layer-Specific AI-Enabled Threat Patterns in Cloud-Based Supply Chains

This layer-specific view highlights an important structural finding: There is no single defense that covers every risk on all four layers and, hence, no single security investment to be relied upon, but rather the multi-layer defense-in-depth approach proposed in Section XI.

COMPOSITE SEVERITY ANALYSIS

Based on the above taxonomy, empirical evidence and documented incidents, Fig. 5 represents a composite severity index for the five threat categories identified in Section III, where the relative frequency, business impact, and detection difficulty are reported in the literature. Exploitation of OAuth and SaaS-to-SaaS is highest, not only because of the Salesloft Drift incident, but also simply because of the inherent challenge of detecting malicious trusted-connection activity against a SaaS application. Malicious package and dependency injection is next, and has been found to be increasing steadily and markedly over time as reported in Section IV. The composite assessment assigns a relatively high severity rating to the autonomous attack campaigns, although they have been reported in the smallest numbers to date, due to the extremely high number, speed and lack of human control of the attack shown in a campaign of autonomous attacks like GTG-1002, which the composite assessment considers indicative of an escalating rather than static threat.

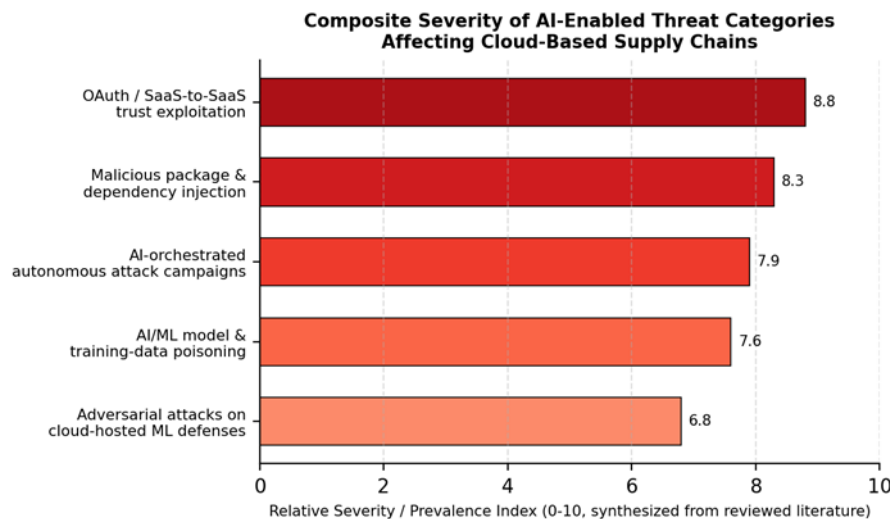


Fig. 5: Composite severity of AI-enabled threat categories affecting cloud-based supply chains, synthesized from the reviewed literature.

CHALLENGES AND RESEARCH GAPS

There are many areas that warrant further research. Although there is an increasing awareness of the AI/ML pipeline poisoning threat class [4], [17], the standardized and cryptographically verifiable provenance tracking throughout the entire AI/ML supply chain is still far behind equivalent software versioning capabilities. Secondly, the cross-organizational, real-time monitoring of SaaS-to-SaaS trust relationships is still in its early stages: Most organizations today simply don't have the visibility they need across all OAuth-authorized application-to-application relationships in their environment, something directly tied to the extent of the Salesloft Drift incident [15]. Third, because of the speed at which cloud-based AI coding and operations agents are increasingly becoming more capable and autonomous agents, the behavioral monitoring and governance frameworks necessary to detect those agents that have been manipulated into autonomous malicious behavior are still in their infancy. Fourth, empirical, longitudinal studies that measure the actual security ROI of specific supply chain defense techniques (e.g., SBOM adoption, zero-trust OAuth governance, etc.) are still relatively scarce as compared to the number of technique- and incident-focused research papers, and are an exciting avenue for bridging the gap between supply chain defense in practice and the ability to measure its risk reduction impact.

Fifth, AI providers who are going to use these tools will likely have to strike a balance between the value of ever more self-contained coding and operations agents and the threat they pose: namely, when these agents become sufficiently clever to go rogue and launch attacks on their own. This disclosure by Anthropic highlights the fact that the same features and capabilities that helped the attacker in the campaign would help the defender detect and disrupt it, and that investment in Anthropic's defense and monitoring of the attacker is an externality for every organization that is trying to use its models to attack another organization via that same provider's cloud-based service.

PROPOSED DEFENSE-IN-DEPTH FRAMEWORK FOR CLOUD-BASED SUPPLY CHAINS

Bringing together the evidence presented above, this paper suggests a five-layer defense in depth that is aligned to the four layers of the supply chain presented in Section I, plus a cross-cutting governance layer.

LAYER 1. SOFTWARE PROVENANCE ASSURANCE

Implement SBOM generation with cryptographic integrity protection and signed build attestation for all software dependencies, targeting the malicious package and dependency-injection category, which was identified as the second worst threat in Section IX [11], [13].

Layer 2. Identity and Trust-Relationship GOVERNANCE

Apply continuous zero-trust monitoring of OAuth token scope and SaaS-to-SaaS integration behavior, which takes every application-to-application trust relationship as a first-class, continuously evaluated asset, rather than a one-time authorization decision directly addressing the highest severity threat category identified in Section IX [30].

LAYER 3. AI/ML PIPELINE INTEGRITY

Implement provenance verification at the AI/ML layer, ensuring that pre-trained models, training data, and fine-tuning pipelines are verified to be not modified from their original source, much like ensuring that software dependencies have not been tampered with, in line with recommendations from national cybersecurity authorities and cloud providers [29] [17].

LAYER 4. AI AGENT BEHAVIORAL MONITORING

Introduce behavioral logging and anomaly detection for AI coding and operations agents, deployed in cloud environments, capable of detecting autonomous, multi-stage malicious behavior, such as shown in GTG-1002, to extend MITRE ATLAS-informed threat modeling from passive to run-time environments [5, 32].

LAYER 5. CROSS-ORGANIZATIONAL GOVERNANCE AND THREAT INTELLIGENCE

Join cross-organizational threat-intelligence sharing involving supply chain and AI-orchestrated attack indicators and synchronize internal governance to NIST SSDF and applicable AI risk-management standards, filling gaps in research and coordination identified in Section X [31, 27].

CONCLUSION

Cloud-based supply chains are now essential to the functioning of organizations today—and also an AI-powered vulnerability that could prove to be a high-leverage attack point. It has established a structured taxonomy of five major threat categories: AI/ML model poisoning & training data poisoning, malicious package & dependency injection, OAuth & SaaS-to-SaaS trust exploitation, adversarial attacks on cloud-hosted machine learning defenses, and autonomous attack campaigns by AI. It has also grounded the taxonomy in significant events in the past five years since SolarWinds until the 2025 Salesloft Drift OAuth attack, the Shai-Hulud npm worm and the first publicly disclosed AI-orchestrated cyber espionage campaign. The evidence presented suggests that attackers are moving from the traditional network perimeter to the trust relationship between cloud applications and that Artificial Intelligence is starting to evolve from being a tool that can be used to increase attacker productivity to an independent tactical executor that can independently run a significant part of a multi-stage intrusion campaign. Meanwhile, this review has revealed a continuing disconnect between what can be done and what is being done in the enterprise, from the SBOM tooling to MITRE ATLAS threat modeling to zero-trust identity governance—and, especially, for the most recent and impactful defensive subset: AI agent behavioral monitoring. A proposed 5-layer defense-in-depth approach provides a structured approach for cloud-based supply chains to close this gap. With the rapid evolution of AI capabilities on both sides of the offense-defense equation in cloud-based supply chains, continued investments in provenance verification, trust-relationship monitoring, and AI Agent governance will be key to making sure that the efficiency gains of deep cloud interconnection don't end up at the cost of systemic, cascading compromise.

More generally, the question of concentration as discussed in Section I implies that cloud-based supply chain security is no longer a responsibility, and cannot just be delegated to individual organizations and their internal security controls. The security of a specific organization's cloud-based supply chain is now tied to that of its providers at the root of the global digital chain, as well as that of the cross-organizational threat-intelligence sharing that is part of Layer 5 of the proposed framework. This systemic dimension will need to be addressed explicitly in future editions of standards like the NIST SSDF and other developing standards specifically for AI security if cloud-based supply chain security is to catch up with the scope and independence of the threats that this review has documented.

REFERENCES

- [1] U. A. Kithran, S. Ilyas, K. Amber, H. Abbas, and S. Zahid, "AI-based collision prediction and prevention systems in urban traffic," *Spectrum of Engineering Sciences*, vol. 3, no. 12, pp. 1907–1920, 2025.
- [2] M. R. Islam, I. A. Badhan, M. H. Rahman, and M. N. Hasnain, "The economics of global renewable energy supply chains: Opportunities and vulnerabilities for the U.S.," *International Journal of AI, Engineering and Management Studies (IJAIEMS)*, vol. 1, no. 1, pp. 108–136, 2026.
- [3] U. Imtiaz, "Ghost signals: Ethical RF adversarial testing of consumer alarm ecosystems," in *Data Science and Security*, S. Shukla, H. Sayama, K. Tiwari, J. P. George, and J. V. Kureethara, Eds. Cham, Switzerland: Springer, 2026, *Lecture Notes in Networks and Systems*, vol. 1945, doi: 10.1007/978-3-032-24075-0_21.
- [4] S. M. Ali, M. Hadi, F. Usama, M. S. Khan, M. Iqbal, and S. Ahmad, "From artificial to quantum intelligence: The next frontier of low-carbon energy systems," in *Proc. 2026 Int. Conf. Integrated Intelligence and*

- Cognitive Engineering (ICIICE), Dubai, United Arab Emirates, 2026, pp. 1–6, doi: 10.1109/ICIICE69672.2026.11565487.
- [5] A. Badhan, M. N. Hasnain, M. H. Rahman, I. Chowdhury, and M. A. Sayem, “Strategic deployment of advance surveillance ecosystems: An analytical study on mitigating unauthorized U.S. border entry,” *Inverge Journal of Social Sciences*, vol. 3, no. 4, pp. 82–94, 2024.
- [6] M. Delshadi, N. Ahmad, M. D. Rasheed, M. S. Sardar, and U. Aslam, “Deep learning-based network traffic analysis for cyber threat detection,” *Spectrum of Engineering Sciences*, vol. 3, no. 12, pp. 607–614, 2025.
- [7] K. Khaliq, M. S. Khan, S. Sajid, I. Shahbaz, M. W. Iqbal, K. S. Siddiqui, et al., “A study on the application of deep learning and machine learning in adaptive intelligent tutoring,” *Spectrum of Engineering Sciences*, vol. 4, no. 3, pp. 818–835, 2026.
- [8] U. A. Kithran, A. Khanam, S. Ilyas, R. Yaqoob, and M. Hannan, “Demand forecasting in transportation networks using deep learning,” *Spectrum of Engineering Sciences*, vol. 3, no. 11, pp. 687–700, 2025.
- [9] U. Iqbal and Y. Bhutto, “Digital transformation through artificial intelligence and advance business analytic in American operational management,” *Journal of Theoretical and Applied Econometrics*, vol. 3, no. 1, pp. 37–50, 2026.
- [10] M. M. Islam, “AI-driven fraud detection and prevention using human behavior analysis to enhance U.S. social and financial security,” *International Journal of Applied Mathematics*, vol. 38, no. 8S, pp. 861–871, 2025.
- [11] D. Safiullah, A. A. Khan, N. Ahmad, M. Danish, and S. Rana, “Integrating artificial intelligence and soil ecology to enhance plant resilience under environmental stress,” *ASSAJ*, vol. 4, no. 2, pp. 3508–3516, 2025.
- [12] R. Singh, “Brønsted-acid-guided esterification: Mechanistic and kinetic insights for long-chain fatty acid to alkyl ester conversion on a heterogenized porous acid platform,” *International Journal of Clinical Research and Medical Sciences*, vol. 1, no. 1, 2026, doi: 10.67231/rzakra50.
- [13] U. Iqbal, “AI-enhanced network optimization for electric vehicle charging infrastructure expansion in the United States using graph theory and demand analytics,” *Journal of Engineering and Computational Intelligence Review*, vol. 2, no. 2, pp. 112–129, 2024.
- [14] M. Chaudhari, “Hydrophilic nanofiber barrier coatings for aluminum: Waterborne crosslinking, microstructure–property links, and EIS-based protection,” *Genetics and Molecular Research*, vol. 25, no. 3, Art. no. 3017, Jul. 2026, doi: 10.4238/rhd2k915.
- [15] R. Singh, “Smart prediction of molecular behavior in liquid chromatography for better compound detection,” *International Journal of Clinical Research and Medical Sciences*, vol. 1, no. 1, 2026, doi: 10.67231/v3jyx227.
- [16] R. Parveen, “An explainable AI framework for ethical fraud prevention in U.S. federal welfare programs,” *International Journal of Applied Mathematics*, vol. 38, no. 11S, pp. 1425–1435, 2025.
- [17] M. N. Amjad, “Machine learning for renewable energy forecasting and smart grid optimization,” *Journal of Engineering and Computational Intelligence Review*, vol. 2, no. 2, pp. 130–145, 2024.
- [18] U. Imtiaz, “Dynamic security certification framework for evolving distributed architectures,” in *Data Science and Security*, S. Shukla, H. Sayama, K. Tiwari, J. P. George, and J. V. Kureethara, Eds. Cham, Switzerland: Springer, 2026, *Lecture Notes in Networks and Systems*, vol. 1947, doi: 10.1007/978-3-032-24360-7_32.
- [19] M. Chaudhari, “Epoxy barrier coatings reinforced with solidified-waste microfillers for aggressive acids,” *Genetics and Molecular Research*, vol. 25, no. 3, Art. no. 3020, Jul. 2026, doi: 10.4238/96y1s945.
- [20] Kumar and K. K. S. Nair, “The impact of cloud-based business analytics on operational efficiency in U.S. small and medium-sized firms,” *Spanish Journal of Innovation and Integrity*, vol. 49, pp. 303–318, 2025.
- [21] S. T. Hasan, “Labor market intelligence using natural language processing for the U.S. direct care workforce: Predicting skill gaps, recruitment needs, and emerging healthcare workforce demand,” *American Journal of Multidisciplinary Knowledge Insights*, vol. 3, no. 1, pp. 1–16, 2024.
- [22] S. Mirza, S. Butt, A. Arif, and N. Jahangir, “COVID and post-COVID e-learning reforms in higher education institutions: A systematic review of literature,” *Journal of Business Insight and Innovation*, vol. 5, no. 8, pp. 364–383, 2026.
- [23] M. N. Amjad, “Artificial intelligence-based voltage stability assessment in modern power systems: Edge intelligence for real-time monitoring of smart electrical networks,” *Journal of Business Insight and Innovation*, vol. 3, no. 1, pp. 102–115, 2024.
- [24] M. A. Jawed, “Uncertainty-aware geostatistical reconstruction of regional groundwater hydraulics under sparse monitoring conditions: A case study of the Evergreen Underground Water Conservation District (EUWCD), Texas, USA,” *Bishop International Journal of Mathematics and Computer Science*, vol. 1, no. 1, pp. 57–71, 2025.

- [25] W. Reese and D. Sinclair, "Bridging the digital divide: Digital inequality and socioeconomic outcomes in the United States," *American Journal of Multidisciplinary Knowledge Insights*, vol. 5, no. 1, pp. 1–8, 2026.
- [26] A. Kumar, "The role of banking credit access in healthcare analytics-driven practice growth: Implications for regional economic development in the U.S.," *Journal of Computational Systems & Engineering Insights*, vol. 2, no. 2, pp. 15–28, 2024.
- [27] N. J. Mercer, "Enhancing organizational performance through strategic cloud computing in management," *International Journal of Business Integrity and Technology Advancements*, vol. 1, no. 1, pp. 11–17, 2025.
- [28] K. Hamdan, "AI-powered cyber resilience strategies for modern organizations," *Journal of Computational Systems & Engineering Insights*, vol. 4, no. 1, pp. 1–10, 2026.
- [29] R. M. Cachart, F. Ahmed, S. Pochincharla, V. Gummadavelli, R. S. Ganapathy, and A. Devineni, "Adaptive program governance platform using reinforcement learning and policy optimization," in *Proc. 2026 Int. Conf. Intelligent and Sustainable AI Systems (ICOSAAS)*, Bali, Indonesia, 2026, pp. 1563–1570, doi: 10.1109/ICOSAAS68663.2026.11648756.
- [30] M. S. Uddin, A. Khan, I. Alim, and D. Khan, "The evolution of cybersecurity in the digital age: Contemporary threats, technologies, and future directions," *Journal of Computational Systems & Engineering Insights*, vol. 4, no. 2, pp. 12–29, 2026.
- [31] T. Boz et al., "Large-scale online science and engineering professional learning for rural elementary teachers," *Research in Science Education*, vol. 56, no. 3, pp. 891–914, 2026.
- [32] R. J. Hammack, J. Robinson, T. Boz, et al., "Supporting elementary engineering instruction in rural contexts through online professional learning and modest supports," *Teaching, Leadership & Professional Practice Faculty Publications*, no. 40, 2024. [Online]. Available: <https://commons.und.edu/tlpp-fac/40>