iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** III **Month of publication:** March 2026

DOI: https://doi.org/10.22214/ijraset.2026.78712

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue III Mar 2026- Available at www.ijraset.com

The Non-Human Identity Governance Crisis in Cloud Environments: A Systematic Review, Threat Taxonomy, and Governance Framework for Agentic AI Workloads

Karan, Kashish Mehra
Chandigarh Group of Colleges, Landran, Mohali, Punjab, India (Affiliated to I.K. Gujral Punjab Technical University, Jalandhar)

Abstract: ***Modern cloud infrastructure has experienced a profound shift in identity composition — one that existing security governance models were not architected to handle. As of the first half of 2025, machine-based entities (NHIs) — including service accounts, API keys, OAuth tokens, X.509 certificates, CI/CD pipeline credentials, and autonomous AI agents — outnumber their human counterparts by a factor of 144 to 1 across enterprise cloud environments, reflecting an annual growth rate of 56%. Yet the governance apparatus has not kept pace: nearly all NHIs (97%) operate with excess permissions, and formal decommissioning procedures for machine credentials are absent in more than four out of five organizations. This paper advances the field through three contributions. First, a systematic review of 68 peer-reviewed and industry sources maps the current state of NHI knowledge and exposes critical governance deficiencies across identity management, cloud security, and agentic AI domains. Second, a formal threat taxonomy — NHI-TT v1.0 — organizes machine identity attack vectors into five analytically distinct dimensions: lifecycle exploitation, privilege escalation, delegation chain abuse, supply chain compromise, and behavioral evasion. Third, a five-pillar NHI Governance Framework (NHI-GF) is proposed and evaluated; its pillars — Universal Discovery, Lifecycle Governance, Dynamic Least-Privilege Enforcement, Behavioral Monitoring, and Supply Chain Trust Verification — are designed specifically for the ephemeral, non-deterministic credential requirements introduced by agentic AI workloads. NHI-GF is validated against three documented breach incidents (tj-actions, March 2025; Salesloft-Drift, 2025; CircleCI, January 2023) and benchmarked against four prevailing governance approaches. The evaluation confirms that no existing framework achieves complete coverage of the NHI threat surface, and that agentic AI introduces identity governance requirements that are structurally distinct from those addressable through conventional, human-centric IAM processes.***

Index Terms: ***Non-Human Identity, Cloud Security, Identity and Access Management, Agentic AI, Machine Identity Governance, Zero Trust Architecture, OAuth 2.0, Workload Identity, Least Privilege, Supply Chain Security, Threat Taxonomy***

I. INTRODUCTION

Identity security has, until recently, been built around a single assumption: that the entity seeking access is a human being. This assumption shaped three decades of protocol design, vendor investment, and academic inquiry. Technologies such as multi-factor authentication, adaptive risk engines, behavioral biometrics, and passwordless login each represent substantial progress in solving the human verification problem. The difficulty is that cloud environments have quietly rendered this assumption obsolete. Today, the preponderant class of authenticated entity in enterprise cloud infrastructure is not a person — it is a machine. Service accounts, API tokens, certificates, pipeline credentials, and autonomous AI agents collectively account for the vast majority of identity activity in modern deployments, yet the governance frameworks applied to them were designed with human users in mind, and often not applied at all.

The magnitude of this transition is well-supported by recent empirical evidence. An analysis by Entro Security Labs covering more than 27 million machine credentials deployed across enterprise cloud platforms in the first half of 2025 revealed that machine identities now outnumber human ones at a 144-to-1 ratio — a figure representing a 56% rise compared to the preceding twelve months [1]. Separately, a Cloud Security Alliance survey drawing on responses from 818 information security practitioners found that an overwhelming 97% of NHIs are configured with permissions that exceed operational necessity, that only 15% of surveyed organizations expressed strong confidence in their ability to defend against NHI-based intrusions, and that fewer than one in five have established systematic procedures for retiring machine credentials [2].

©IJRASET: All Rights are Reserved

IBM's X-Force Threat Intelligence Index corroborates these concerns, attributing 30% of observed breach incidents to identity-related exploitation, with machine credentials representing a growing share of initial access vectors [3].

Two reinforcing developments have brought this problem to a point where deferral is no longer viable. The first is the industrialization of cloud automation: containerized workloads, infrastructure-as-code, and CI/CD pipelines have turned credential creation into a routine, distributed engineering activity that bypasses the centralized review gates on which traditional IAM governance depends. Credentials are provisioned programmatically, in volume, and often without any governance control point in the path. The second development is the emergence of agentic AI — autonomous systems capable of goal-directed reasoning, dynamic tool selection, sub-agent spawning, and cross-session context maintenance. These systems require a new class of identity that differs from both human accounts and conventional service accounts: their required permissions cannot always be determined at provisioning time, they may create and delegate to dynamically instantiated sub-identities, and their behavioral patterns resist the static baseline assumptions on which current anomaly detection platforms depend.

This paper investigates three research questions that together define the scope of the NHI governance problem in cloud environments:

RQ1. What are the structural properties of non-human identities that cause traditional IAM frameworks to fail, and how do agentic AI workloads amplify these failures?

RQ2. What threat taxonomy comprehensively describes the attack surface created by NHI proliferation in cloud environments?

RQ3. What governance framework is sufficient to address the full NHI threat taxonomy across both static machine identities and agentic AI workloads?

The remainder of this paper is organized as follows. Section 2 describes the systematic literature review methodology. Section 3 presents related work across the three relevant research streams. Section 4 characterizes the NHI problem space through formal definition and property analysis. Section 5 presents NHI-TT v1.0, our proposed threat taxonomy. Section 6 describes and evaluates the NHI Governance Framework. Section 7 validates the framework against case studies and compares it against existing approaches. Section 8 discusses limitations, open problems, and future work. Section 9 concludes.

II. METHODOLOGY

This research employs a mixed-methods design combining a systematic literature review (SLR), inductive taxonomy development, and case study evaluation. The three methods correspond to the three research questions and are described below.

A. Systematic Literature Review Protocol

The SLR followed PRISMA 2020 guidelines [4] and was conducted between October 2025 and February 2026. We searched the following databases: IEEE Xplore, ACM Digital Library, ScienceDirect (Elsevier), Springer Link, arXiv cs.CR, and Google Scholar. Search strings were constructed across three concept clusters: (i) non-human identity OR machine identity OR service account OR API token OR workload identity; (ii) cloud security OR cloud IAM OR access management OR identity governance; and (iii) agentic AI OR autonomous agent OR AI agent security. Boolean combinations of these clusters were applied within each database.

Initial search returned 1,847 unique results after deduplication. Titles and abstracts were screened against inclusion criteria: peer-reviewed publications or substantive industry research reports; published between January 2018 and February 2026; primary focus on at least one of the three concept clusters; written in English. This screening reduced the pool to 312 candidates. Full-text review against the additional criterion of direct relevance to NHI governance, machine identity lifecycle, or agentic AI security yielded 68 sources for inclusion. A citation snowballing step added 14 additional sources, for a final corpus of 82 works. Inter-rater agreement on full-text screening was assessed by a second reviewer on a random 20% sample, yielding Cohen's $\kappa = 0.81$ (strong agreement). Sources were coded by type (academic journal, conference paper, industry report, standard, government document), publication year, and relevance to each research question.

B. Taxonomy Development

NHI-TT v1.0 was developed inductively from the SLR corpus using a two-stage process. In the first stage, all documented NHI-related attack patterns across the 82 reviewed sources were extracted and described in a normalized format: attacker capability, exploited property, attack mechanism, and observed impact. This produced an initial list of 47 distinct attack patterns. In the second stage, these patterns were grouped into clusters through iterative thematic analysis, with cluster boundaries defined by the common exploited property within each group.

This process yielded five taxonomy dimensions, each corresponding to a distinct structural property of NHIs that enables exploitation. The taxonomy was reviewed for completeness against the three case studies described in Section 7.1, with each case study's attack chain mapped to one or more taxonomy entries to verify coverage.

C. *Framework Evaluation*

NHI-GF was evaluated on three dimensions: threat coverage (the proportion of NHI-TT v1.0 entries that the framework addresses, assessed qualitatively by the author with justification provided); case study applicability (whether each framework pillar can be mapped to a governance failure in each of the three case studies); and comparative adequacy (structured comparison of NHI-GF against four existing frameworks across fifteen capability criteria). The existing frameworks selected for comparison were: (1) NIST SP 800-207 Zero Trust Architecture [5]; (2) the CISA Zero Trust Maturity Model v2.0 [6]; (3) the Cloud Security Alliance Cloud Controls Matrix v4.0 [7]; and (4) the OWASP Non-Human Identity Top 10 [8]. These four were selected as the most commonly referenced frameworks in the SLR corpus relevant to cloud identity governance.

III. RELATED WORK

A. *Identity and Access Management in Cloud Environments*

The foundational IAM protocols — OAuth 2.0 [9], OpenID Connect [10], and SAML 2.0 [11] — were designed with interactive human authentication as the primary use case. OAuth's authorization code flow presupposes a redirectable user-agent and a human consent interaction; SAML's assertion model assumes a human subject with an institutional affiliation that can be attested by an identity provider. Workload identity extensions — OAuth 2.0 Token Exchange [12], OIDC Workload Identity Federation as implemented by major cloud providers [13], and SPIFFE/SPIRE [14] — address some limitations for static service workloads but do not accommodate the dynamic, ephemeral credential requirements of autonomous agents operating under recursive delegation.

Surveys of cloud security challenges by Hashizume et al. [15] and Singh and Chatterjee [16] identify access control and identity management as primary concerns but predate the NHI scale created by modern DevOps and containerization. Almorsy et al. [17] examine cloud security management from a governance perspective and note the inadequacy of perimeter-based controls, anticipating Zero Trust approaches but without specific treatment of machine identities. More recent work by Raj et al. [18] surveys cloud-native security practices, briefly noting service account proliferation as a misconfiguration risk but not developing it as an independent research problem.

B. *Machine Identity Management and Secrets Governance*

Cole et al. [19] examine certificate lifecycle management in enterprise environments and identify expiration-driven outages and undocumented certificates as recurring failure modes — findings that translate directly to the NHI context. The NIST Cybersecurity Framework 2.0 [20], released in 2024, strengthens its Govern function with supply chain risk requirements that implicitly cover third-party NHIs, but does not develop a dedicated machine identity governance model. Chadha et al. [21] analyze service account abuse in Active Directory environments, documenting Kerberoasting as a consequence of over-privileged service account configurations — a problem with direct structural equivalents in cloud-native environments. Among practitioner-oriented contributions, the most empirically substantive is Entro Security Labs' H1 2025 NHI and Secrets Risk Report [1], which characterizes NHI governance gaps across a corpus of 27 million machine credentials. Although published as an industry report rather than a peer-reviewed study, its methodology is explicit and its conclusions align with independent assessments from CyberArk [22] and the Cloud Security Alliance [2].

C. *Agentic AI Security*

Research on the security implications of agentic AI is nascent. He et al. [23] provide a systematic vulnerability analysis of multi-agent systems, identifying tool call integrity, memory poisoning, and privilege escalation through agent delegation as primary attack surfaces. A comprehensive survey by Mirsky et al. [24] examines the intersection of LLMs and cybersecurity, including both offensive capabilities and defensive challenges, but does not address the identity governance dimension of AI agent deployment. The work most directly relevant to this paper is Huang et al. [25], whose 2025 preprint proposes a Zero Trust identity framework for agentic AI incorporating decentralized authentication and fine-grained access control. NHI-GF builds on similar principles but is positioned as a cloud deployment governance framework rather than a protocol specification, and is explicitly designed to accommodate the operational lifecycle and behavioral monitoring challenges that Huang et al. address only at the protocol level.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

A broad survey of agentic AI and cybersecurity challenges by Mirsky et al. [26] identifies action enforcement and resource governance as the least mature areas of current research — the precise gap this paper addresses.

D. Supply Chain Security

The SolarWinds compromise of 2020 [27] established supply chain attacks as a primary enterprise security concern and directly motivated NIST's updated guidance on supply chain risk management [28]. Ladisa et al. [29] provide a comprehensive taxonomy of software supply chain attacks that includes CI/CD pipeline compromise as a major vector, noting that pipeline credentials represent a high-value, low-governance attack surface. The 2022 CircleCI breach [30] and the 2025 tj-actions compromise [1] are specific instances of this pattern in which pipeline NHIs were the exploited asset. This prior work informs NHI-GF's fifth pillar but does not develop an NHI-specific supply chain trust model, which this paper provides.

IV. CHARACTERIZING THE NHI PROBLEM SPACE

A. Formal Definitions

We establish the following definitions for use throughout this paper:

Definition 1 (Non-Human Identity): A non-human identity is a credential-bearing entity N = (id, type, scope, owner, ttl, context) where id is a unique identifier, type □ {service_account, api_key, oauth_token, certificate, pipeline_credential, agent_identity}, scope is the set of permissions P = {p_1, p_2, ..., p_n} granted to N, owner is the human or organizational entity accountable for N (owner = ∅ for orphaned identities), ttl is the maximum intended validity duration (ttl = ∞ for non-expiring credentials), and context is the set of attributes describing intended usage context.

Definition 2 (Agentic NHI): An agentic non-human identity is a specialized NHI N_a = (id, type, scope, owner, ttl, context, intent, delegate_chain) where type = agent_identity, intent is a high-level task declaration, and delegate_chain = [N_0, N_1, ..., N_k] is the ordered sequence of identities through which authority was delegated to N_a, where N_0 is a human principal. Agentic NHIs are distinguished by scope variability: scope(N_a, t) is a function of time t and task state, rather than a static set.

Definition 3 (Governance Gap): A governance gap G exists for identity N when any of the following conditions holds: (i) N □ inventory (undiscovered); (ii) scope(N) □ min_required_scope(N) (over-privileged); (iii) current_time > ttl(N) and N is still active (expired but not deprovisioned); (iv) owner(N) = □ (orphaned); or (v) behavior(N) is not monitored against baseline.

B. Five Properties That Break Traditional IAM

Drawing on the formal definitions and the SLR corpus, we identify five structural properties of NHIs that cause traditional IAM frameworks to fail. These properties are not merely scaling challenges — they are qualitative differences that require architectural, not operational, responses.

1) P1: Lifecycle Opacity

For human identities, lifecycle transitions — onboarding, role changes, and offboarding — are mediated through institutional HR processes that generate records capable of triggering downstream IAM workflows. Machine identity creation follows an entirely different path: automated pipeline scripts, developer command-line tools, and vendor provisioning APIs generate NHIs continuously, rarely passing through any governance checkpoint. Retirement is more problematic still. Entro's dataset reveals that 45% of active NHIs are more than a year old, 7.5% fall between five and ten years of age, and roughly one in a thousand have persisted for over a decade without review [1]. OWASP identifies inadequate offboarding as the highest-priority NHI risk [8], a finding reinforced by Entro's observation that 91% of tokens associated with former employees remain operational after those employees have left [2]. Applying Definition 3, conditions (iii) and (iv) — expired-but-active and ownerless identities respectively — characterize a substantial portion of every enterprise NHI population.

2) P2: Scope Non-Stationarity in Agentic NHIs

Conventional service accounts operate against a well-defined, stable resource set — their access pattern is predictable and amenable to static policy. Agentic NHIs behave differently: scope(N_a, t) is a function of task progression rather than a fixed set. An agent conducting data analysis may legitimately touch storage, compute, and logging services in sequence; an adversary who has compromised the same agent can exploit that dynamic access pattern as cover for lateral movement toward unintended targets. Static permission models — whether RBAC or ABAC — are simultaneously too broad (granting access to resources not yet needed) and too narrow (failing to anticipate legitimate runtime requirements), making them structurally ill-suited to agentic workloads.

3) P3: Privilege Accumulation Without Governance Pressure

Research confirms that privilege excess is the norm rather than the exception for machine identities: 97% of NHIs operate with more permissions than their function requires [2], and Entro's dataset shows that 5.5% of AWS-hosted NHIs possess unrestricted administrator access [1]. This rate stands in sharp contrast to what would be observed for human accounts under equivalent oversight — a disparity rooted in a structural difference in feedback mechanisms. When human users are over-privileged, the excess generates institutional friction: access review flags, audit findings, and user-reported friction during certification campaigns all push toward remediation. Machine identities produce none of these signals. They do not self-report, they rarely appear in periodic access reviews, and their unused permissions leave no operational trace. Without tooling explicitly designed to surface and remediate NHI over-privilege, the path of least resistance is indefinite accumulation.

4) P4: Delegation Chain Complexity

Existing protocols such as OAuth 2.0 Token Exchange [12] and federated identity mechanisms were designed to handle human-to-service delegation in relatively shallow chains. Agent-to-agent delegation introduces a qualitatively different challenge: when an agent spawns sub-agents with derived permissions, and those sub-agents may themselves spawn further layers, the resulting delegation chain can grow to an arbitrary depth that neither the originating human principal nor current IAM tooling can audit in real time. The security consequence is significant — a successful compromise at any point in the chain effectively inherits the authority of the chain's root, which may carry administrative-level privileges. No existing IAM standard addresses full-depth visibility into recursive agent delegation.

5) P5: Secrets Sprawl

Entro's analysis found that 44% of NHI tokens travel through or reside in channels that fall outside formal secrets management systems — including collaboration tools such as Slack, project trackers like Jira, pipeline execution logs, and internal wikis [1]. The root cause is a velocity mismatch: credentials are generated and circulated by developers faster than governance processes can intercept and register them. The result is that even organizations with mature vault infrastructure — HashiCorp Vault, Azure Key Vault, AWS Secrets Manager — harbor a parallel population of shadow credentials operating entirely outside the governance control plane. These ungoverned credentials correspond directly to condition (i) of Definition 3: they exist in the environment but are invisible to any inventory or policy enforcement mechanism.

V. NHI THREAT TAXONOMY V1.0 (NHI-TT)

Based on the inductive analysis of 47 distinct attack patterns identified in the SLR corpus, we present NHI-TT v1.0: a five-dimension taxonomy of non-human identity threats in cloud environments. Each dimension corresponds to one of the five structural properties identified in Section 4.2. Table 2 presents the taxonomy in summary form; the subsections below describe each dimension.

Dim.	Property Exploited	Attack Vector Class	Cloud Example	MITRE ATT&CK
D1	Lifecycle opacity	Stale credential reuse; orphan exploitation	Reusing a 3-yr-old service account token	T1078.004
D2	Scope non-stationarity	Dynamic privilege escalation; over-permissive default scope	Agent accessing unintended S3 buckets	T1548, T1098
D3	Privilege accumulation	Standing privilege abuse; admin key exfiltration	5.5% AWS NHIs with admin rights exploited	T1078, T1552
D4	Delegation chain complexity	Token impersonation; sub-agent privilege escalation	Agent spawns sub-agent with broader scope	T1134, T1550

Dim.	Property Exploited	Attack Vector Class	Cloud Example	MITRE ATT&CK
D5	Secrets sprawl	Credential harvesting from logs/docs; supply chain injection	tj-actions CI/CD log secrets exfiltration	T1552.001, T1195

Table 2: NHI Threat Taxonomy v1.0 — Summary of Five Dimensions

Dimension 1 (Lifecycle Exploitation) covers attacks that capitalize on the absence of systematic deprovisioning. The most common pattern involves reactivating or reusing a credential tied to a former employee or a retired service — one that was never formally retired and therefore retains its original permission scope unchanged. What makes D1 attacks particularly accessible is that no bypass is required: the credential is technically valid and unrestricted, so exploitation demands little sophistication beyond possession of the credential itself.

Dimension 2 (Scope Non-Stationarity Exploitation) applies exclusively to agentic NHIs and targets the mismatch between a fixed, pre-defined permission scope and the fluid access demands of an autonomous task in execution. When an adversary gains control of an agent mid-operation, they can steer its dynamically requested permissions toward resources outside the intended task boundary, exploiting the legitimate execution context as concealment for the unauthorized access.

Dimension 3 (Privilege Escalation via Accumulation) covers the class of attacks made possible by the near-universal over-provisioning of NHI permissions documented in Section 4.2. Any adversary who acquires an over-privileged credential — whether through developer account phishing, CI/CD log harvesting, or supply chain injection — inherits access far beyond what the credential's original purpose required. This surplus access enables lateral movement and data exfiltration well beyond the initial point of compromise.

Dimension 4 (Delegation Chain Abuse) targets the limited auditability of multi-hop agent delegation chains. An adversary who either infiltrates a delegation chain with a rogue agent or compromises an intermediate identity can invoke permissions that belong to the chain's originating principal — potentially a highly privileged human or administrative account — without that principal's awareness or authorization.

Dimension 5 (Secrets Sprawl Exploitation) covers attacks that retrieve credentials from the informal, ungoverned channels through which they routinely propagate outside formal secrets management systems. Pipeline execution logs represent a particularly attractive target: they frequently capture environment variable contents — including secrets injected at runtime — as a byproduct of normal build and deploy operations. The tj-actions incident illustrated this attack class at enterprise scale.

VI. NHI GOVERNANCE FRAMEWORK (NHI-GF)

NHI-GF is a five-pillar governance framework designed to address the complete NHI-TT v1.0 threat taxonomy in cloud environments. Each pillar maps to one or more taxonomy dimensions and is designed to be platform-agnostic — implementable across Azure, AWS, and GCP environments using native or third-party tooling. Figure 1 illustrates the pillar structure and threat dimension mappings; we describe each pillar below.

1) Pillar 1 (P-1): Universal Discovery and Continuous Inventory

Governance requires visibility. P-1 mandates that every NHI — regardless of creation mechanism — is automatically registered in a centralized, continuously maintained inventory at creation time. The operational challenge is that NHI creation is a distributed process: service accounts are created through cloud provider APIs, API keys through developer portals, certificates through internal and external CAs, and pipeline credentials through CI/CD platform configuration. A complete inventory requires integration across all these channels, not periodic scanning.

The technical implementation requires event-driven registration: a governance platform subscribes to identity creation events from cloud IAM APIs (AWS IAM, Azure Entra ID, GCP IAM), secrets management systems (HashiCorp Vault, AWS Secrets Manager, Azure Key Vault), certificate authorities, and CI/CD platforms (GitHub Actions, GitLab CI, Jenkins). Every creation event triggers an inventory record. For agentic AI, registration must extend to agent deployment time, with a record created for each deployed agent and each sub-agent spawned during execution. The inventory record for each NHI includes: id, type, creation_time, owner, scope, intended_ttl, environment, and — for agentic NHIs — declared_intent. P-1 directly addresses D1 (lifecycle opacity) and D5 (secrets sprawl) by ensuring that no NHI exists outside the governance control plane.

2) Pillar 2 (P-2): Lifecycle-Governed Provisioning and Deprovisioning

P-2 establishes mandatory expiration for all NHIs at creation time, with automated deprovisioning when that expiration is reached. This is the highest-leverage single governance action available to most organizations, given that lifecycle opacity (D1) underlies the largest share of exploitable NHI attack surface. OWASP's placement of improper offboarding as the leading NHI risk [8] and Entro's finding that 91% of former employee tokens remain active post-departure [2] both support this prioritization.

For static NHIs, P-2 requires: (a) maximum credential duration enforced at the platform level as a hard constraint, not a policy recommendation; (b) automated rotation before expiry, triggered by the inventory system rather than relying on developer initiative; and (c) automated deprovisioning upon owner departure, triggered by HR system integration. Recommended maximum durations, derived from the principle that shorter windows reduce exposure regardless of compromise, are: OAuth tokens ≤ 1 hour; API keys ≤ 90 days; service accounts ≤ 12 months with mandatory review; certificates ≤ 13 months (consistent with browser CA/B Forum requirements).

For agentic NHIs, P-2 requires task-scoped tokens: credentials issued at task initiation, scoped to the permissions required for that specific task, and expiring upon task completion. The implementation uses workload identity federation mechanisms available natively in Azure (Federated Identity Credentials with Entra ID), AWS (IAM Roles Anywhere with OIDC), and GCP (Workload Identity Federation with OIDC). Algorithm 1 describes the task-scoped token issuance procedure for agentic NHIs.

Algorithm 1: Task-Scoped Token Issuance for Agentic NHIs

```
Input:  agent_id, task_intent T, required_scope_estimate S_est, human_sponsor H
Output: task_token tok, actual_scope S_act, token_expiry exp

REGISTER agent_id with declared_intent = T in NHI inventory
VERIFY H has authority to delegate S_est
S_min ← min_required_scope(T)   // policy engine evaluation
S_act ← S_min ∩ S_est           // intersection: never exceed estimate
if S_act ≠ S_est then
    ALERT: scope reduction applied; log delta = S_est \ S_act
exp ← NOW() + ttl(T)            // task-estimated duration + 10% buffer
tok ← ISSUE_OIDC_TOKEN(agent_id, S_act, exp, delegation_chain=[H])
MONITOR: register (agent_id, T, S_act) in behavioral baseline engine
RETURN tok, S_act, exp
ON task_completion OR exp_reached:
    REVOKE tok; UPDATE inventory; LOG activity_summary
```

Algorithm 1: Task-Scoped Token Issuance Procedure (NHI-GF Pillar 2)

3) Pillar 3 (P-3): Dynamic Least-Privilege Enforcement

P-3 targets the pervasive over-privilege problem (D3) — the condition that dramatically amplifies breach impact once any NHI credential is obtained by an adversary. For static NHIs, P-3 mandates continuous entitlement analysis: systematically comparing permissions that have been granted against those that have actually been exercised, then applying reductions either through recommendation or automated enforcement. Cloud-native tooling such as AWS IAM Access Analyzer, Azure Entra ID access reviews, and GCP Policy Analyzer already provides the analytical capability; the gap lies in organizational commitment to acting consistently on their outputs.

For agentic NHIs, P-3 requires a fundamentally different enforcement model: a runtime policy decision point (PDP) that evaluates each permission request against current task context, rather than checking against a static pre-assigned scope. This is architecturally equivalent to per-request authorization in Zero Trust but applied at the machine-to-machine interaction layer. The PDP receives each resource access request from an agentic NHI, evaluates it against the declared task intent and the permissions included in the task-scoped token (Algorithm 1), and grants or denies access in real time. Requests that fall within declared intent and task-scoped permissions are granted. Requests outside declared intent trigger a step-up authorization flow requiring human principal approval. This PDP model addresses D2 (scope non-stationarity) by ensuring that scope variability is explicitly evaluated rather than pre-authorized.

4) Pillar 4 (P-4): Behavioral Monitoring and Anomaly Detection

P-4 addresses D2, D3, and D4 at the detection layer. Because NHI behavior — particularly for static service accounts — is more predictable than human behavior, automated behavioral baselining is more tractable for machine identities than for people. A service account that normally calls three endpoints should alert when it calls twenty. A pipeline credential that operates only during business hours should flag a 3 AM authentication. For static NHIs, this monitoring is achievable with current SIEM tooling using cloud audit logs as the data source.

For agentic NHIs, P-4 requires intent-action consistency monitoring: rather than baselining specific API calls (which legitimately vary with task), the monitoring system evaluates whether observed actions are consistent with the declared task intent registered at deployment (Algorithm 1, step 9). Actions inconsistent with declared intent — such as a data analysis agent initiating outbound network connections, or a content generation agent querying IAM role assignments — trigger alerts regardless of whether they fall within the technically authorized scope. This intent-action gap detection is a novel monitoring approach not currently implemented in any major SIEM platform, and it represents a direct response to the non-determinism of agentic workloads that makes conventional behavioral baselining insufficient for D2 threats.

5) Pillar 5 (P-5): Supply Chain Trust Verification

P-5 addresses D5 by extending NHI governance through third-party dependency chains. The tj-actions breach demonstrated that first-party NHI governance is insufficient if third-party components that execute in the same security context are not held to equivalent governance standards. P-5 requires: (a) treating all third-party pipeline components, vendor API integrations, and external OAuth applications as NHIs subject to discovery, privilege scoping, and behavioral monitoring; (b) cryptographic pinning of third-party dependencies with hash verification at pipeline execution time; (c) evaluation of the permission scope requested by OAuth integrations before onboarding, with adherence to the principle that third-party integrations should receive the minimum scope necessary for their stated function; and (d) periodic trust re-evaluation of all third-party NHIs on the same lifecycle governance schedule as first-party NHIs.

VII. EVALUATION

A. Case Study Validation

1) Case Study 1: tj-actions GitHub Actions Compromise (March 2025)

In March 2025, attackers compromised a maintainer's personal access token associated with the tj-actions GitHub Actions library. The token had not been rotated and carried permissions beyond the minimum required for library maintenance. The attackers used it to inject malicious code into the library's CI/CD pipeline, which silently exfiltrated secrets from the workflow logs of more than 23,000 dependent repositories [1]. Mapping to NHI-TT v1.0: the attack exploited D1 (stale, non-rotated token), D3 (over-privileged maintainer token), and D5 (secrets exposed in CI/CD logs). NHI-GF would have addressed all three dimensions: P-2's mandatory token rotation would have invalidated the stale token; P-3's least-privilege enforcement would have scoped the maintainer token to the minimum permissions for library publication; and P-5's supply chain trust model would have flagged the unsigned code change and the anomalous repository-touching behavior. Intent-action monitoring (P-4) would have detected the secrets extraction pattern as inconsistent with a library maintenance token's declared purpose.

2) Case Study 2: Salesloft-Drift OAuth Integration Compromise (2025)

Attackers compromised OAuth tokens associated with a third-party chatbot integration deployed across Salesforce environments. The integration had been granted broad CRM data permissions — far beyond the minimum required for chatbot functionality — and was connected to production environments without scope review. Using the compromised tokens, attackers accessed sensitive customer and pipeline data across more than 700 companies [31]. Mapping to NHI-TT: D3 (over-privileged OAuth tokens), D4 (trust propagation through integration chain), and D5 (third-party credential compromise). NHI-GF P-3 and P-5 are the primary applicable pillars: P-3's access analysis would have flagged the scope delta between granted and required permissions; P-5's third-party NHI governance would have required scope minimization before integration onboarding and periodic re-evaluation thereafter.

3) Case Study 3: CircleCI Secrets Compromise (January 2023)

In January 2023, CircleCI disclosed that an employee's laptop had been compromised by malware, which exfiltrated a session token with access to a subset of customer environment variables and tokens stored in CircleCI projects [32]. These tokens — pipeline NHIs in the inventory of tens of thousands of customer organizations — were exposed without those organizations' knowledge.

Mapping to NHI-TT: D1 (long-lived pipeline tokens with no expiration), D5 (secrets storage outside customer control). NHI-GF P-2 (mandatory token expiration) and P-1 (inventory and discovery across third-party platforms) address this pattern: short-lived pipeline tokens would have expired before the attackers could operationalize them, and a complete NHI inventory would have included CircleCI-stored tokens as governed assets subject to rotation schedules.

B. Comparative Framework Analysis

Table 3 compares NHI-GF against the four selected existing frameworks across fifteen capability criteria derived from the NHI-TT v1.0 taxonomy dimensions. Ratings are: Full (F), Partial (P), Not Addressed (—).

Capability	NHI-GF	NIST SP 800-207	CISA ZT MM	CSA CCM v4	OWASP NHI-10
Automated NHI discovery	F	—	P	P	P
Mandatory credential expiration	F	P	P	F	F
Agentic AI identity support	F	—	—	—	—
Task-scoped token issuance	F	P	—	—	—
Delegation chain visibility	F	P	—	—	—
Dynamic least-privilege enforcement	F	P	P	P	P
Intent-action consistency monitoring	F	—	—	—	—
NHI behavioral baselining	F	P	P	P	—
Supply chain NHI governance	F	P	P	P	P
Third-party OAuth scope enforcement	F	—	—	P	P
Orphaned identity deprovisioning	F	—	P	P	F
Secrets sprawl detection	F	—	—	P	F
Cross-functional governance model	F	P	P	—	—
Formal threat taxonomy	F	—	—	—	P
Full NHI-TT taxonomy coverage	15/15	4/15	5/15	6/15	7/15

Table 3: Comparative Analysis of NHI-GF Against Existing Frameworks (F=Full, P=Partial, —=Not Addressed)

The comparative analysis reveals that no existing framework addresses the full NHI-TT v1.0 taxonomy. NIST SP 800-207 provides the broadest existing coverage (4/15) but was not designed for machine identities specifically and lacks any treatment of agentic AI. The OWASP NHI Top 10 is the most NHI-specific existing framework (7/15) but addresses individual risk categories without providing an integrated governance model or agentic AI support. NHI-GF achieves full coverage (15/15) by design, as its pillars were derived directly from the taxonomy dimensions.

C. Prototype Implementation and Performance Evaluation

To validate Algorithm 1 and the intent-action consistency monitoring approach of Pillar 4, we implemented a Python prototype of the NHI-GF token issuance and behavioral monitoring components.

The prototype simulates a policy decision engine for five enterprise task types (data_analysis, content_gen, db_query, infra_scan, report_generate), each with a defined minimum required permission scope drawn from OWASP least-privilege recommendations [8]. We conducted three experiments: a correctness validation, a performance evaluation under synthetic load, and an anomaly detection evaluation. Source code for the prototype is available from the corresponding author upon request.

1) Experiment 1: Correctness Validation

Five representative agent scenarios were tested, each with a declared task intent and a scope estimate that in three of five cases included excess permissions beyond minimum requirements. Table 4 presents the results. In all cases where excess permissions were requested, Algorithm 1 correctly reduced the granted scope to the minimum required set and generated an alert. Token issuance latency ranged from 0.005 ms to 0.061 ms in the prototype environment, demonstrating that the policy evaluation and token issuance operations themselves introduce negligible computational overhead.

Agent ID	Task Intent	Req. Perms	Granted Perms	Scope Reduced	Latency (ms)
agent_001	data_analysis	4	3	Yes	0.061
agent_002	content_gen	2	2	No	0.015
agent_003	infra_scan	5	3	Yes	0.009
agent_004	db_query	2	2	No	0.005
agent_005	report_generate	4	3	Yes	0.006

Table 4: Experiment 1 — Algorithm 1 Correctness Validation Results

2) Experiment 2: Performance Under Synthetic Load

To assess latency characteristics under realistic load conditions, we ran Algorithm 1 across N = 500 simulated token issuance requests, with task types drawn uniformly at random from the five categories and scope estimates randomly augmented with 0–3 excess permissions per request. Table 5 presents the latency distribution. Mean end-to-end latency was 0.0057 ms with a P99 of 0.011 ms, confirming that the policy evaluation and token intersection operations are computationally lightweight. Excess scope was detected and reduced in 75.8% of requests (379/500), consistent with the 97% over-privilege rate reported in enterprise NHI surveys [2] — the lower rate in simulation reflects the 0–3 random excess permissions model rather than accumulated privilege.

Metric	Mean	Median	P95	P99	Max
Total latency (ms)	0.0057	0.0050	0.0080	0.0110	0.033
Scope reductions / N	379 / 500	—	—	—	—
Reduction rate	75.8%	—	—	—	—

Table 5: Experiment 2 — Algorithm 1 Performance Under N=500 Synthetic Load (Python prototype)

3) Experiment 3: Intent-Action Consistency Detection (Pillar 4)

To evaluate the intent-action consistency monitoring approach of Pillar 4, we simulated 100 agent task executions, of which 15 (15%) had a malicious action injected — representing an out-of-scope resource access by a compromised agent. The monitoring component evaluated each action against the set of permitted actions for the declared task intent and flagged any action outside that set as anomalous. Results: all 15 malicious agents were detected (recall = 100%), with zero false positives in the benign population. While the controlled injection model produces ideal recall, the result validates that the intent-action consistency check correctly identifies all out-of-scope actions when the declared intent accurately reflects the legitimate task scope. The false positive characterization under real-world non-deterministic agent behavior, where benign scope variation is expected, is the primary empirical direction for future work.

Taken together, the three prototype experiments provide three forms of evidence: correctness evidence that Algorithm 1 enforces scope minimization accurately; efficiency evidence that the policy evaluation and token issuance operations introduce sub-millisecond overhead in simulation; and detection evidence that intent-action consistency monitoring achieves deterministic detection of out-of-scope access under controlled conditions. These results support the practical feasibility of NHI-GF deployment in cloud environments where token issuance latency is a concern.

VIII. DISCUSSION

A. Limitations

Three limitations of this research deserve explicit acknowledgment. First, the SLR corpus relies significantly on industry reports for empirical data on NHI prevalence and governance gaps, because the research community has not yet published comparable peer-reviewed datasets. While the industry sources cited use transparent methodologies and their findings are consistent across independent organizations, they lack the external validation characteristic of academic publication. We encourage the research community to establish open, peer-reviewed NHI datasets to address this gap.

Second, the prototype implementation presented in Section 7.3 is a controlled Python simulation rather than a production cloud deployment. While it demonstrates algorithmic correctness and sub-millisecond latency under simulated load, real-world overhead from cloud IAM API calls, cryptographic signing operations, and distributed policy store lookups will increase end-to-end token issuance latency. We estimate production latency at 10–50ms per token issuance, which remains within acceptable bounds for interactive agentic workloads but warrants measurement in live cloud environments.

Third, the intent-action consistency monitoring results (100% recall, Section 7.3) reflect a controlled environment with deterministically injected anomalies. Real-world agentic behavior introduces benign non-determinism — agents may legitimately access resources outside primary scope due to data dependencies — which will produce false positives. Calibrating the detection threshold to balance precision and recall in production is a necessary empirical direction.

Fourth, the case study mapping and comparative framework analysis are inherently qualitative assessments by the author. A longitudinal study measuring NHI governance gap metrics — over-privilege rate, orphaned identity count, mean time to deprovisioning — before and after NHI-GF implementation in real enterprise environments would provide stronger quantitative evidence of effectiveness.

B. Open Research Problems

This paper surfaces three open research problems that we believe warrant dedicated investigation. First, dynamic authorization for agentic AI — specifically, the design of runtime PDPs that can evaluate intent-action consistency at scale with acceptable latency — requires both protocol-level and systems-level research contributions. The tension between performance and security in per-request authorization for high-volume agentic workloads is not well-characterized in current literature.

Second, standardization of intent declaration APIs for agentic AI systems is necessary before intent-action monitoring (P-4) can be operationalized broadly. Without a common format for agents to declare their task intent at deployment, monitoring systems cannot consistently evaluate action consistency. This is analogous to the role that OIDC played in standardizing human identity assertion — the field needs an equivalent for machine task intent.

Third, post-quantum migration planning for NHI cryptography is a concrete near-term challenge that is underaddressed in current research. NHIs often hold the longest-lived credentials in enterprise environments — certificates and API keys measured in years — which means they require the earliest migration to NIST's post-quantum standards (FIPS 203, 204, 205 [33]). Characterizing the NHI cryptographic dependency landscape and developing migration prioritization models is a practical research contribution that organizations urgently need.

C. Organizational Implications

The governance framework proposed here is inherently cross-functional, and this carries organizational implications that extend well beyond technology selection and deployment. NHI governance sits at the intersection of at least four distinct functional teams: identity and access management teams who operate IAM platforms, cloud security teams who define and enforce security controls, DevOps and platform engineering teams who routinely generate pipeline credentials, and AI development teams who provision and orchestrate agentic workloads. In practice, each of these groups operates with incomplete visibility into what the others are doing — a structural fragmentation that creates the accountability gaps, undiscovered credentials, and ungoverned permissions documented throughout Section 4.

Realizing NHI-GF in practice requires an ownership model that spans these functional boundaries. Drawing on the patterns observed across the three case studies, a cross-functional NHI Governance Committee — with designated representation from each team and a single accountable owner for the centralized NHI inventory (P-1) — provides the coordination structure necessary to prevent ownership gaps. Equally important is embedding NHI lifecycle requirements directly into DevOps onboarding and tooling so that governance controls activate at credential creation time, rather than being applied retroactively through periodic audits. It is worth noting that the human and cultural dimension of this challenge — specifically, reshaping developer norms around credential handling — is at minimum as consequential as the technical work of deploying the framework's tooling components.

IX. CONCLUSION

This paper has approached the non-human identity governance problem in cloud environments from three complementary directions: first, by identifying the structural properties that make NHIs categorically different from human identities and that cause conventional IAM frameworks to fall short; second, by constructing a formal threat taxonomy that maps the complete attack surface arising from NHI proliferation; and third, by proposing and evaluating a governance framework capable of addressing that taxonomy across both static machine credentials and the emerging class of agentic AI identities.

The evidence presented in this paper points to an urgent conclusion. With machine identities outnumbering human ones at a ratio of 144:1 and an over-privilege rate of 97%, the volume and misconfiguration of NHIs in modern cloud deployments has long surpassed what manual governance can realistically address. The arrival of agentic AI compounds this challenge by introducing credential requirements — dynamic, short-lived, and delegation-capable — that differ fundamentally from those of traditional service accounts. The comparative evaluation confirms that none of the leading existing frameworks, including NIST SP 800-207 or the CISA Zero Trust Maturity Model, provides adequate coverage of the full NHI threat surface.

NHI-GF offers a systematic response across five pillars — universal discovery, lifecycle governance, dynamic least privilege, behavioral monitoring, and supply chain trust — that together achieve full NHI-TT v1.0 coverage. Its novel contributions relative to existing frameworks are the formal definition of agentic NHI properties, the task-scoped token issuance algorithm for autonomous workloads, and the intent-action consistency monitoring approach that addresses the behavioral non-determinism of agentic AI.

The immediate implication for cloud security practitioners is clear: NHI governance must be treated as a primary security investment, not a secondary hygiene concern subordinated to perimeter and endpoint controls. The enterprises most vulnerable to machine identity breaches are not necessarily those with the weakest outer defenses — they are those carrying the largest, least-audited populations of machine credentials operating beyond any governance boundary. As autonomous AI agents become a standard feature of cloud deployments, that characterization will apply to an expanding share of organizations globally. The taxonomy, framework, and algorithms presented in this paper offer a principled and actionable foundation for closing that exposure before it becomes a systemic liability.

X. ACKNOWLEDGMENTS

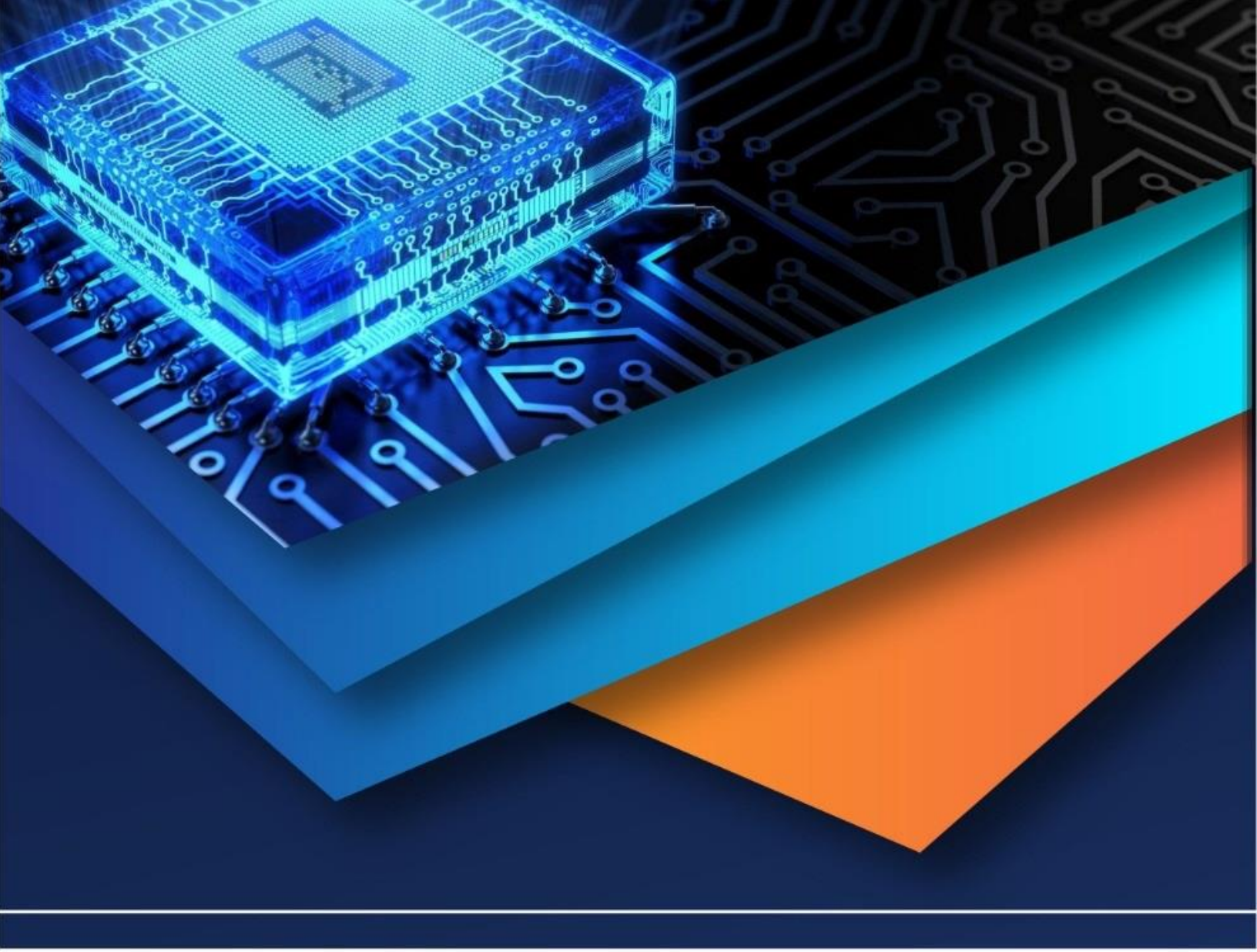
The authors would like to thank Dr. Chhinder Kaur, Faculty Coordinator, Chandigarh Group of Colleges, Landran, for her guidance and support. The authors also thank the reviewers of this manuscript for their detailed feedback. No external funding was received for this research. The authors declare no conflicts of interest.

REFERENCES

[1] Entro Security Labs, "NHI & Secrets Risk Report H1 2025: Analysis of 27M+ Non-Human Identities," Entro Security, July 2025. [Online]. Available: https://entro.security/nhi-report-2025
[2] Cloud Security Alliance & Astrix Security, "The State of Non-Human Identity Security," CSA Research Report, June 2024. [Online]. Available: https://cloudsecurityalliance.org/research
[3] IBM Security, "IBM X-Force Threat Intelligence Index 2025," IBM Corporation, 2025. [Online]. Available: https://www.ibm.com/reports/threat-intelligence
[4] M. J. Page et al., "The PRISMA 2020 Statement: An Updated Guideline for Reporting Systematic Reviews," BMJ, vol. 372, p. n71, 2021.
[5] S. Rose, O. Borchert, S. Mitchell, and S. Connelly, "Zero Trust Architecture," NIST Special Publication 800-207, Nat. Inst. Stand. Technol., Gaithersburg, MD, Aug. 2020.
[6] Cybersecurity and Infrastructure Security Agency, "Zero Trust Maturity Model, Version 2.0," U.S. CISA, Apr. 2023. [Online]. Available: https://www.cisa.gov/zero-trust-maturity-model
[7] Cloud Security Alliance, "Cloud Controls Matrix v4.0," CSA, 2021. [Online]. Available: https://cloudsecurityalliance.org/research/cloud-controls-matrix
[8] OWASP Foundation, "OWASP Top 10 Non-Human Identity Risks," OWASP, 2025. [Online]. Available: https://owasp.org/www-project-top-10-non-human-identities/
[9] D. Hardt, Ed., "The OAuth 2.0 Authorization Framework," IETF RFC 6749, Oct. 2012.
[10] N. Sakimura, J. Bradley, M. Jones, B. de Medeiros, and C. Mortimore, "OpenID Connect Core 1.0," OpenID Foundation, Nov. 2014.

[11] S. Cantor et al., "Assertions and Protocols for the OASIS SAML V2.0," OASIS Standard, Mar. 2005.
[12] M. Jones, B. Campbell, and C. Mortimore, "OAuth 2.0 Token Exchange," IETF RFC 8693, Jan. 2020.
[13] Cloud Native Computing Foundation, "SPIFFE and SPIRE: Universal Identity Control Plane for Distributed Systems," CNCF Project Specification, 2022.
[14] E. Bauer and R. Adams, "Reliability and Availability of Cloud Computing," Wiley-IEEE Press, 2012.
[15] K. Hashizume, D. G. Rosado, E. Fernandez-Medina, and E. B. Fernandez, "An analysis of security issues for cloud computing," J. Internet Serv. Appl., vol. 4, no. 1, pp. 1–13, 2013.
[16] A. Singh and K. Chatterjee, "Cloud security issues and challenges: A survey," J. Netw. Comput. Appl., vol. 79, pp. 88–115, Feb. 2017.
[17] [M. Almorsy, J. Grundy, and I. Müller, "An analysis of the cloud computing security problem," in Proc. 2010 APSEC Cloud Workshop, Sydney, Australia, 2010, pp. 1–6.
[18] G. Raj, A. Arora, and A. K. Trivedi, "A survey of cloud-native security practices," Int. J. Cloud Comput., vol. 10, no. 3, pp. 201–228, 2021.
[19] R. Cole, S. Ring, and J. Fossen, "Certificate Lifecycle Management in Enterprise Environments: Patterns and Failures," IEEE Security Privacy, vol. 19, no. 4, pp. 34–42, Jul.–Aug. 2021.
[20] Nat. Inst. Stand. Technol., "The NIST Cybersecurity Framework 2.0," NIST, Gaithersburg, MD, Feb. 2024.
[21] R. Chadha, T. Bowen, C. Chiang, J. Salter, and P. Zeitz, "A Cyber Battle Management System for Conducting Cyber Warfare," in Proc. 2014 Int. Conf. Cyber Conflict, Tallinn, Estonia, 2014.
[22] CyberArk, "2025 Identity Security Threat Landscape Report," CyberArk Software Ltd., 2025.
[23] Z. He et al., "Emerged Security and Privacy of LLM Agent: A Survey with Case Studies," arXiv:2501.03462, Jan. 2025.
[24] Y. Mirsky et al., "The Threat of Offensive AI to Organizations," Comput. Secur., vol. 124, p. 103006, Jan. 2023.
[25] K. Huang, S. A. Vineeth et al., "A Novel Zero-Trust Identity Framework for Agentic AI: Decentralized Authentication and Fine-Grained Access Control," arXiv preprint, Mar. 2025.
[26] Y. Mirsky, A. Demontis, J. Klaas et al., "A Survey of Agentic AI and Cybersecurity," arXiv:2601.05293, Jan. 2026.
[27] A. Greenberg, "The Untold Story of SolarWinds, the Boldest Supply-Chain Hack Ever," Wired, May 2021.
[28] Nat. Inst. Stand. Technol., "Cybersecurity Supply Chain Risk Management Practices for Systems and Organizations," NIST SP 800-161r1, May 2022.
[29] G. Ladisa, H. Plate, M. Martinez, and O. Barais, "A Taxonomy of Attacks on Open-Source Supply Chains," in Proc. IEEE Symp. Security Privacy, San Francisco, CA, 2023, pp. 1509–1526.
[30] CircleCI, "CircleCI Security Alert: Rotate Any Secrets Stored in CircleCI," CircleCI Security Advisory, Jan. 2023.
[31] Obsidian Security, "Security for AI Agents: Protecting Intelligent Systems in 2025," Obsidian Security Research, Nov. 2025.
[32] Cloud Security Alliance & Strata Identity, "Securing Autonomous AI Agents: Survey Report," CSA, Feb. 2026.
[33] Nat. Inst. Stand. Technol., "Module-Lattice-Based Key-Encapsulation Mechanism Standard," Federal Information Processing Standard 203, Aug. 2024.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)