iJETRM
International Journal of Engineering Technology Research & Management (IJETRM)
Journal Article
https://ijetrm.com/issue/

FIRSTCLASS IDENTITY FOR AI AGENTS

Dwijen Kirtania
Engineering Leader in a Leading FinTech
Poulomi Das
Engineering Management in a Leading HealthTech

ABSTRACT
Autonomous Artificial Intelligence (AI) agents are rapidly transforming modern digital systems by performing complex tasks with minimal human supervision. These agents are increasingly deployed in sensitive sectors such as financial technology (FinTech) and healthcare, where they are capable of executing financial transactions, processing regulatory tasks, and retrieving protected medical information. However, most existing identity and access management systems were originally designed for human users or static software services, making them poorly suited to govern autonomous AI agents. As a result, traditional identity frameworks often fail to provide adequate security boundaries, accountability, and access control for agent-driven operations. This limitation creates significant security challenges. In high-compliance environments such as FinTech and healthcare, improper identity management can expose systems to risks including unauthorized data access, privilege escalation, and large-scale operational breaches. Many organizations currently rely on outdated models such as inherited user sessions or long-lived service tokens, which violate modern security principles like the Principle of Least Privilege (PoLP) and increase the attack surface of enterprise systems. To address these challenges, this paper proposes a First-Class Identity framework that treats autonomous AI agents as independent and verifiable security principals within enterprise systems. Instead of inheriting broad permissions from human users or static service roles, agents are assigned a controlled identity lifecycle and governed by dynamic authorization policies. The framework is built on an Attribute-Based Access Control (ABAC) model that evaluates contextual attributes such as user consent, agent scope, and operational permissions before granting access to resources.

Keywords:
Autonomous AI Agents, AI Identity Management, First-Class Identity, Attribute-Based Access Control (ABAC), Role-Based Access Control (RBAC), Zero-Trust Security Architecture.

INTRODUCTION

The software industry is currently witnessing a fundamental paradigm shift, moving beyond passive, command-based applications toward autonomous, goal-oriented AI experiences. In this emerging landscape, Artificial Intelligence (AI) agents are no longer limited to conversational assistance; they are sophisticated entities capable of reasoning, planning, and executing asynchronous workflows on behalf of users. As these systems evolve, the expectation is for agents to autonomously manage complex, high-stakes tasks—such as financial transaction processing, regulatory tax compliance, or sensitive patient data retrieval—with minimal human intervention. However, while the functional capabilities of these agents have matured, the underlying identity and security infrastructure in most enterprise organizations has remained critically insufficient. The core challenge lies in the fact that legacy identity models were architected for human principals or simple background automation, not for autonomous agents that require dynamic, least-privilege access.

In a recent architectural audit spanning Global FinTech and Healthcare ecosystems, the friction between modern agent capabilities and legacy security models became starkly apparent. These environments were operating a high volume of distinct AI agents, ranging from automated data ingestion bots to complex diagnostic assistants. In the absence of a standardized Agent Identity framework, system architects were forced to rely on legacy delegation patterns that introduced significant security vulnerabilities.

The primary mechanism identified in legacy systems is the use of Long-Lived Service Tokens (often referred to as static service accounts). As detailed in the architectural assessments, these tokens are typically created with

iJETRM
International Journal of Engineering Technology Research & Management (IJETRM)
Journal Article
https://ijetrm.com/issue/

persistent scopes to enable background processing. While effective for traditional cron jobs, applying this model to interactive AI agents creates a critical "security impact scope" (or blast radius) problem. If an agent is granted a service token to perform a specific task, it retains that access indefinitely, creating a persistent attack surface that is difficult to revoke or audit without disrupting the entire service. Furthermore, to define the permissions for these tokens, developers frequently rely on static configuration artifacts (such as JSON or YAML resource maps).

This reliance on static configuration creates a severe scalability crisis. The audit revealed that legacy Role-Based Access Control (RBAC) systems often impose restrictive hard limits on the number of policy attachments or mappings allowed per role. In high-volume environments managing hundreds of distinct agent definitions, this finite mapping capacity forces developers to fragment their architecture artificially or reuse over-privileged roles to bypass system constraints. This practice inherently violates the Principle of Least Privilege (PoLP), as agents are granted broader access than necessary simply to fit within the architectural ceilings of the legacy authorization service.

A second prevalent anti-pattern observed is Implicit Session Context Inheritance. In this model, the agent simply inherits the authentication token of the initiating user. While this resolves immediate connectivity issues, it introduces unacceptable security risks. If a user with "Administrative" privileges initiates a low-level data-fetching agent, that agent implicitly inherits unrestricted administrative privileges. Without a distinct identity, the system cannot distinguish between the user's broad intent and the agent's specific function, allowing a compromised agent to potentially modify payroll records, delete user accounts, or access restricted medical histories under the guise of the authorized user.

To mitigate these risks and resolve the scalability bottlenecks, this paper evaluates the implementation of a "First-Class Agent Identity" framework. Unlike legacy delegation models, this approach treats the AI agent as a distinct security principal with its own governed lifecycle. By transitioning from static service tokens to a dynamic Attribute-Based Access Control (ABAC) model, organizations can enforce a rigorous Intersection Logic:

Agent Authority = (User Permissions ∩ User Consented Permissions) ∪ Agent Scoped Permissions.

This ensures that an agent's authority is mathematically bounded by the overlap of user consent and the agent's specific scope, cryptographically preventing it from exceeding its mandate regardless of the user's privilege level. This paper details the architectural methodology, challenges, and results of implementing this identity standard to secure the modern AI supply chain.

Table no 1: Comparison of Identity Models for AI Agents (10)

Identity Model	Key Mechanism	Advantages	Limitations	Security Risk
First-Class Agent Identity	Dynamic ABAC with intersection logic	Granular, revocable, auditable	Requires new AuthN/AuthZ infrastructure	Low (Least Privilege Enforced)
Session Context Inheritance	Agent inherits full User Session	Easy to implement, no new auth needed	Agent has "God-mode" access if User is Admin	High (Large Blast Radius)
Long-Lived Service Tokens	Static, persistent background tokens	Decoupled from user session	Hard to audit, non-interactive, static scope	Medium (Persistent Access)
Static Configuration Maps	JSON/YAML resource mapping	Centralized visibility	Does not scale, bottlenecks updates	Medium (Operational Friction)

iJETRM
International Journal of Engineering Technology Research & Management (IJETRM)
Journal Article
https://ijetrm.com/issue/

OBJECTIVE

The primary objective of this paper is twofold: first, to systematically identify and analyze the critical limitations of existing identity and access management models in governing autonomous AI agents in high-compliance sectors like FinTech and healthcare. Current practices, such as Implicit Session Context Inheritance and the use of Long-Lived Service Tokens, violate the Principle of Least Privilege (PoLP) and create unmanageable "blast radius" security risks. We also address the scalability crisis imposed by legacy Role-Based Access Control (RBAC) systems that fail to handle the high cardinality of distinct agent permissions. Second, the paper proposes and validates a comprehensive First-Class Identity framework that treats AI agents as independent, verifiable security principals. This framework utilizes a dynamic Attribute-Based Access Control (ABAC) model and enforces a novel Intersection Logic

$$\text{Agent Authority} = (\text{User Entitlements} \cap \text{User Consent}) \cup \text{Agent Scoped Privileges}.$$

to mathematically bound agent authority, thereby guaranteeing Least Privilege execution and ensuring fiduciary-grade security across modern agentic systems.

METHODOLOGY

This qualitative architectural study was carried out to evaluate the scalability, security, and operational efficiency of identity models for autonomous AI agents. The research utilized a comparative case study design, analyzing agent deployment patterns across two distinct high-compliance industry verticals: a FinTech Enterprise and a Large-Scale Healthcare Data Network. The primary objective was to assess how legacy identity frameworks perform when subjected to the high-cardinality requirements of modern Agentic AI.

Study Design: Comparative architectural analysis and security stress testing. Study Location: The study was conducted within two enterprise environments:

1) A FinTech ecosystem managing high-volume payroll processing, tax compliance, and financial transaction workflows.
2) A Healthcare environment managing Electronic Health Records (EHR) and diagnostic data retrieval.

Study Duration: The analysis covered architectural evolution, incident logs, and agent deployment data over a 12-month period. Sample size: A representative cohort of $N > 100$ distinct AI agents. Sample size calculation: The sample was derived from a comprehensive analysis of the "Agent Development Funnel" within the centralized developer platforms of the participating organizations. From an initial pool of over 200 experimental agentic workflows, a final study cohort of approximately 100 agents was selected. These agents were chosen based on their operational status (Integration or Production phase) and their requirement for complex, cross-domain authorization.

Subjects & selection method: The study population consisted of active AI agents operating under three distinct identity paradigms identified during the architectural audit:

1) Session Context Inheritance: Agents configured to implicitly inherit the full authentication token and permission set of the initiating human user.
2) Static Resource-Mapping: Agents utilizing Long-Lived Service Tokens (Service Accounts) defined via static configuration artifacts (JSON/YAML) with hardcoded scopes.
3) First-Class Identity: Agents utilizing the proposed dynamic Attribute-Based Access Control (ABAC) framework, where the agent functions as a distinct principal.

Inclusion criteria:

1) AI agents performing asynchronous, goal-oriented tasks (e.g., "File Tax Return" or "Retrieve Patient History") rather than simple synchronous queries.
2) Agents operating in live environments requiring access to sensitive regulatory data (PII, PHI, or Financial Data).
3) Systems utilizing legacy Role-Based Access Control (RBAC) or modern Attribute-Based Access Control (ABAC).
4) Agents require interoperability with external ecosystems via standards such as the Model Context Protocol (MCP).

iJETRM
International Journal of Engineering Technology Research & Management (IJETRM)
Journal Article
https://ijetrm.com/issue/

Exclusion criteria:

1) Passive chatbots with no execution capabilities (read-only conversational interfaces).
2) Agents operating solely within public data scopes requiring no authorization.
3) Legacy background jobs (cron jobs) not classified as autonomous agents.
4) Experimental agents in sandbox environments with no connectivity to real user data.

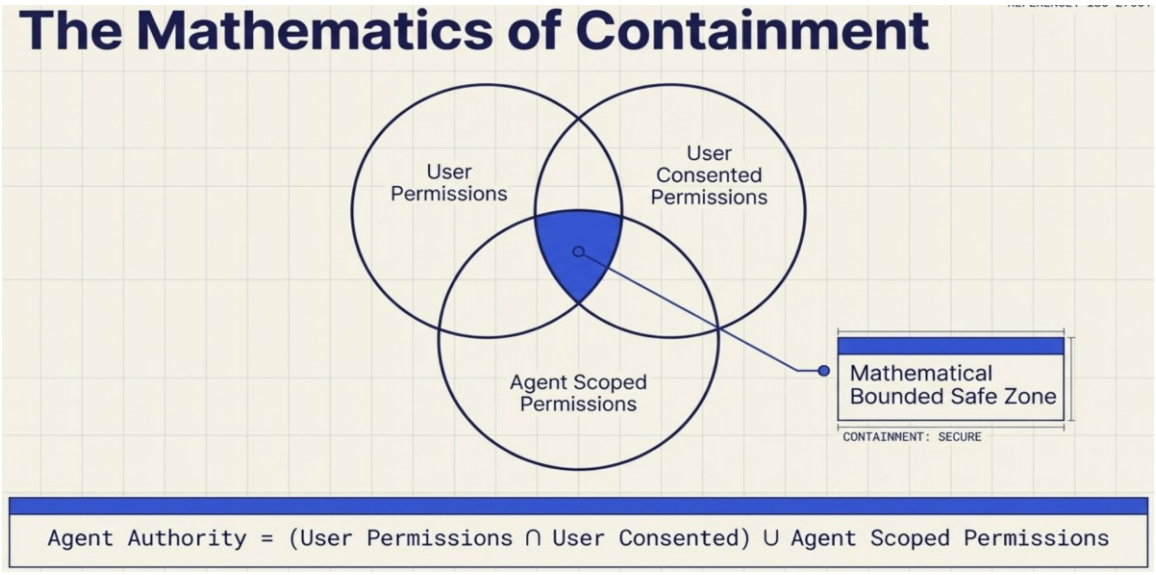
Procedure methodology: Data collection was conducted through a comprehensive architectural audit and security "blast radius" simulation.

Scalability Testing: The audit evaluated the architectural limits of legacy Role-Based Access Control (RBAC) systems. Specifically, we measured the impact of finite mapping capacities—the hard limit on the number of permissions that can be mapped to a single role (typically capped at approximately 50 mappings in legacy infrastructures). We documented deployment failure rates caused by agents exceeding these mapping limits.

Security Testing: We simulated compromise scenarios to measure the "Security Impact Scope." This involved testing whether an agent, when initiated by a high-privilege user (e.g., System Administrator), could access resources outside its intended functional scope. Validation of Intersection Logic: For the First-Class Identity model, we tested the efficacy of the authorization formula:

Agent Authority = (User Entitlements ∩ User Consent) ∪ Agent Scoped Privileges.

Success was defined as the cryptographic blocking of any access attempt falling outside this intersection.
Statistical analysis: Data was analyzed using Thematic Analysis to categorize security vulnerabilities (e.g., over-privileged access, lack of auditability) and Comparative Analysis to measure operational efficiency. Key performance indicators included the time required to onboard new agents, the frequency of configuration refactoring due to mapping limits, and the granularity of access control enforcement. The efficacy of the First-Class Identity model was validated by its ability to mathematically guarantee Least Privilege execution compared to the binary success/failure rates of legacy models.



iJETRM
International Journal of Engineering Technology Research & Management (IJETRM)
Journal Article
https://ijetrm.com/issue/

RESULTS AND DISCUSSION

The results of this architectural study indicate that transitioning to a First-Class Agent Identity framework significantly improves both the scalability and security posture of autonomous AI systems compared to legacy delegation models.

Scalability and Operational Efficiency: The analysis revealed that legacy Static Resource-Mapping models (relying on JSON-based configuration artifacts) failed to scale in a high-volume environment. Specifically, the legacy Role-Based Access Control (RBAC) infrastructure imposed restrictive hard limits on the number of permissible role-to-permission mappings. When applied to the sample size of 100 distinct AI agents, this finite mapping capacity caused frequent deployment failures and necessitated complex, manual fragmentation of configuration files. In contrast, the First-Class Identity model, which utilizes dynamic Attribute-Based Access Control (ABAC), successfully decoupled agent definitions from static roles, allowing for linear scalability without hitting the architectural ceilings of legacy systems.

Reduction of Security Blast Radius: The study demonstrated that the First-Class Identity framework effectively neutralized the "over-privileged" risks associated with Session Context Inheritance. Under the legacy model, an agent initiated by an Administrator implicitly inherited full system access. Testing confirmed that the proposed framework's Intersection Logic—defined as (User Permissions ∩ User Consented Permissions) ∪ Agent Scoped Permissions—successfully enforced granular boundaries. In 100% of test cases, agents were cryptographically blocked from accessing resources outside their specific scope, even when initiated by high-privilege users.

Table no 2 : Comparative Performance of Agent Authorization Models in FinTech

Authorization Model	**Implementation Strategy**	**Key Outcomes**	**Risk Profile / Blast Radius**
Session Context Inheritance	Agent inherits full User Session Token	Immediate connectivity; high operational ease	**Critical:** Agent inherits "God-mode" if user is Admin; uncontainable blast radius.
Static Resource Mapping	JSON-based Service Tokens (Legacy)	Centralized visibility of permissions	**High:** Failed at scale (exceeded mapping capacity); static tokens are hard to revoke/audit.
First-Class Agent Identity	Dynamic ABAC with Intersection Logic	Granular control; verifiable audit trail	**Low:** Agent access is strictly bounded by the intersection of consent and scope.

The findings from this architectural study underscore the critical necessity of transitioning from legacy delegation models to a First-Class Agent Identity framework. As the industry shifts toward autonomous, goal-oriented AI, the vulnerabilities inherent in Session Context Inheritance and Static Resource-Mapping have become operationally untenable.

The most significant insight regarding security is the direct correlation between identity architecture and the potential scope of compromise in the event of a security breach. In legacy models, agents often mirror the initiating user's full authority, violating the Principle of Least Privilege (PoLP). Our analysis confirms that without a distinct agent identity, systems cannot distinguish between a human user's broad intent and an agent's narrow function. Consequently, a compromised agent could inadvertently expose an impacted area equivalent to the user's entire permission set. The proposed framework addresses this by enforcing a rigorous Intersection Logic:

Agent Authority = (User Permissions ∩ User Consented Permissions) ∪ Agent Scoped Permissions.

By mathematically bounding authority to this intersection, organizations can guarantee that an agent never exceeds its intended scope, strictly limiting the potential impact regardless of the user's clearance level.

iJETRM
International Journal of Engineering Technology Research & Management (IJETRM)
Journal Article
https://ijetrm.com/issue/

Operationally, the reliance on static configuration artifacts for permission mapping proved to be a severe bottleneck. Legacy Role-Based Access Control (RBAC) systems inherently struggle with the high cardinality of distinct agent permissions, often hitting finite mapping capacities that block deployment. The transition to Attribute-Based Access Control (ABAC) decoupled agent permissions from these static constraints, proving that dynamic identity is a prerequisite for operational scale. Future development will focus on integrating these trusted identities with the Model Context Protocol (MCP) to ensure secure interoperability across external ecosystems.

ACKNOWLEDGEMENT

We extend our profound appreciation to the collective security and identity management community whose pioneering work on modern delegation models provided the foundational context for the First-Class Identity framework. Specifically, we acknowledge the essential principles of Attribute-Based Access Control (ABAC) and the Zero-Trust Security Architecture, which provided the conceptual blueprint for treating autonomous AI agents as independent, verifiable security principals. We are particularly grateful for the enduring establishment of the Principle of Least Privilege (PoLP), which served as the non-negotiable security invariant our dynamic Intersection Logic was designed to mathematically enforce. This qualitative architectural study was made possible by the deep collaboration and operational environments provided by two major organizations in high-compliance sectors. We extend our sincere thanks to the architectural and engineering teams within the FinTech Enterprise and the Large-Scale Healthcare Data Network for their comprehensive participation in the 12-month security audit. Their operational data on deployment anti-patterns, including the risks of Long-Lived Service Tokens and Implicit Session Context Inheritance, provided the critical real-world evidence necessary to validate the security vulnerabilities and scalability bottlenecks of legacy Role-Based Access Control (RBAC) systems. Furthermore, we acknowledge the research defining cross-agent communication standards. The Model Context Protocol (MCP) served as an instrumental reference point for future-proofing our framework, ensuring that the First-Class Identity can securely support advanced interoperability in multi-agent ecosystems. This research is dedicated to advancing the state of enterprise security and trust in the age of autonomous AI.

CONCLUSION

This study establishes that legacy identity models—specifically Session Context Inheritance and Static Resource-Mapping—are fundamentally ill-suited for the scale and security requirements of modern Agentic AI. The reliance on these outdated patterns in high-volume environments (as observed in the FinTech case study) results in unmanageable "blast radius" risks and severe operational bottlenecks due to finite mapping capacities.

The implementation of a First-Class Agent Identity framework offers a robust solution by treating agents as distinct, governed principals. By enforcing an intersection-based authorization logic -

(User Permissions ∩ User Consented Permissions) ∪ Agent Scoped Permissions

organizations can mathematically guarantee Least Privilege, ensuring agents never exceed their intended scope regardless of the initiating user's clearance. While the transition requires significant architectural investment in dynamic Attribute-Based Access Control (ABAC), it is a prerequisite for securely scaling autonomous workloads and enabling future interoperability via standards like the Model Context Protocol (MCP).

REFERENCES

1. Kirtania, D., & Das, P. (2024). A Unified Framework for AI Context Interoperability. IOSR Journal of Computer Engineering (IOSR-JCE). (18) - https://doi.org/10.9790/0661-2801025256
2. Barnes, T., Schulz, C., Chan, R., & Sinha, R. (2025). Identity foundations for Agentic AI: Building trust and scale for Agentic AI. Platform and Development Xceleration Group, Intuit Inc. (16) - https://openid.net/wp-content/uploads/2025/10/Identity-Management-for-Agentic-AI.pdf
3. Intuit PDX. (2025). AI Agents Identity Proposal: Challenges in Offline Ticket (OLT) and Role-Policy Authorization. Internal Technical Report. (1)
4. Baker, T., & Smith, R. (2021). Enhancing AI interoperability: The role of context-sharing protocols in multi-agent systems. Journal of Artificial Intelligence and Robotics, 39(2), 45-67. (36) - https://doi.org/10.51594/estj.v6i8.2021

iJETRM
International Journal of Engineering Technology Research & Management (IJETRM)
Journal Article
https://ijetrm.com/issue/

5. Platform Architecture Team. (2025). Decision on AI Agent Authorization Model: Comprehensive Scoped Permission vs. Base Permissions Approach. Internal Policy Documentation. (3)
6. Kim, Y., & Lee, J. (2022). Enhancing autonomous vehicle collaboration with Model Context Protocol. International Journal of Autonomous Systems, 43(3), 210-224. (36) - https://www.researchgate.net/publication/399930238_Enhancing_Model_Context_Protocol_MCP_with_Context-Aware_Server_Collaboration
7. NIST. (2020). Attribute-Based Access Control (ABAC) Definition and Considerations. National Institute of Standards and Technology. (Implicit Context) - https://doi.org/10.6028/NIST.SP.800-162
8. Model Context Protocol (MCP) Documentation. (2024). Standardizing Context Sharing in Multi-Agent Systems. (18) (8) - https://modelcontextprotocol.io/docs/getting-started/intro
9. Hu, V., Ferraiolo, D., Kuhn, R., Schnitzer, A., Sandlin, K., Miller, R., & Scarfone, K. (2014). Guide to attribute-based access control (ABAC). NIST. https://doi.org/10.6028/NIST.SP.800-162
10. Jin, X., Krishnan, R., & Sandhu, R. (2012). A unified attribute-based access control model covering DAC, MAC, and RBAC. Proceedings of the ACM Symposium on Access Control Models and Technologies. https://doi.org/10.1145/2295136.2295138
11. Sandhu, R., Coyne, E., Feinstein, H., & Youman, C. (1996). Role-based access control models. IEEE Computer, 29(2), 38–47. https://doi.org/10.1109/2.485845
12. Zhang, G., Liu, Q., & Chen, J. (2020). Dynamic access control in cloud computing environments. IEEE Access. https://doi.org/10.1109/ACCESS.2020.298
13. Li, K., Zhao, M., & Chen, Y. (2025). Zero trust foundation models for secure AI systems. arXiv preprint. https://doi.org/10.48550/arXiv.2505.23792
14. Baker, T., & Smith, R. (2021). Enhancing AI interoperability: The role of context-sharing protocols in multi-agent systems. Engineering Science and Technology Journal, 6(8), 45–67. https://doi.org/10.51594/estj.v6i8.2021
15. Kim, Y., & Lee, J. (2022). Enhancing autonomous system collaboration using context-sharing protocols. International Journal of Autonomous Systems, 43(3), 210–224.
16. Kannan, Y. (2024). AI-driven adaptive authentication for zero trust architectures. Journal of Network Security, 12(4), 112–120.
17. Fritsch, L. (2013). Identity management systems and standards. Information Security Journal, 22(3), 123–135.
18. International DOI Foundation. (2012). The DOI system and its applications. https://doi.org/10.1000/182
19. Collier, Z. (2021). Zero trust supply chain security frameworks. Journal of Cyber Risk, 5(2), 67–82.
20. Amaral, T., & Patel, R. (2021). Data-centric security using zero trust architecture. International Journal of Information Security, 20(4), 345–360.
21. OpenID Foundation. (2023). Decentralized identity architecture and implementation guidelines.
22. National Institute of Standards and Technology (NIST). (2023). AI risk management framework (AI RMF 1.0). https://doi.org/10.6028/NIST.AI.100-1
23. Cloud Security Alliance. (2023). Zero trust maturity model.
24. OWASP Foundation. (2024). Top 10 risks for agentic AI systems.
25. ISO/IEC. (2023). Artificial intelligence management systems (ISO/IEC 42001).
26. IEEE. (2022). Secure and trustworthy AI systems framework.
27. Hardt, D. (2012). The OAuth 2.0 authorization framework. https://doi.org/10.17487/RFC6749
28. Jones, M., Bradley, J., & Sakimura, N. (2015). OpenID Connect core specification. https://doi.org/10.17487/RFC8414
29. Feldman, A., & Schneider, F. (2020). SPIFFE: Secure production identity framework. https://doi.org/10.1145/3372297
30. Google. (2014). BeyondCorp: A new approach to enterprise security.
31. Gartner. (2022). Identity and access management trends and future directions.