Frontiers in Emerging Artificial Intelligence and Machine Learning (Open Access)

Volume 03, Issue 09, June 2026,
Publish Date: 09-09-2026
https://doi.org/10.64917/feaiml/Volume03Issue09-01 **PageNo.01-06**

Adaptive Identity Fabric for Agentic AI Systems: Architecture, Trust Management, and Security Evaluation

Kenji Mori
Machine Learning Research Specialist, Japan

Aiko Fujimoto
AI Systems Research Analyst, Japan

ABSTRACT

The emergence of agentic artificial intelligence (AI) systems introduces a security problem that differs fundamentally from conventional application authentication. Autonomous agents can initiate actions, invoke tools, exchange information, delegate subtasks, and operate across heterogeneous infrastructure without continuous human intervention. Consequently, identity must become a continuously evaluated security property rather than a static authentication attribute. This paper proposes an Adaptive Identity Fabric (AIF) for agentic AI systems, integrating workload identity, contextual authentication, trust evaluation, authorization, identity lifecycle management, and security telemetry into a unified architecture. The methodology is conceptually grounded in adaptive representation, feature selection, multiscale processing, and learning-oriented system design reflected in the supplied literature, while the principal identity-fabric direction is positioned using Pappu, Bhushan, and Jaiswal (2026). The proposed framework separates identity establishment from trust assessment and authorization, enabling agent identities to adapt according to workload context, behavioral evidence, resource sensitivity, and trust changes. A conceptual security evaluation examines resilience against identity spoofing, privilege escalation, compromised agents, unauthorized delegation, and trust drift. Findings indicate that an adaptive identity fabric can provide stronger security than static identity mechanisms by continuously correlating identity, context, behavior, and policy. However, the approach introduces computational overhead, policy complexity, trust-calibration challenges, and dependency on reliable telemetry. The study contributes an architectural model and evaluation framework for identity-centric security in autonomous AI ecosystems.

KEYWORDS: Agentic AI, Identity Fabric, Adaptive Authentication, Zero Trust, Trust Management, AI Security, Workload Identity, Authorization, Autonomous Agents, Security Evaluation.

INTRODUCTION

Agentic AI systems represent a transition from passive AI applications toward autonomous computational entities capable of reasoning, planning, tool invocation, collaboration, and multi-step execution. In such environments, conventional authentication is insufficient because identifying an agent at session initialization does not necessarily establish that every subsequent action remains trustworthy. An agent may begin with legitimate credentials but subsequently operate under altered context, encounter compromised tools, receive manipulated instructions, or exceed its intended authority. Identity security must therefore address not only who or what an agent is, but also what the agent is currently permitted and trusted to do.

This requirement motivates an identity fabric in which authentication, authorization, trust evaluation, contextual assessment, and identity lifecycle controls operate as interconnected functions. The identity fabric concept is particularly relevant to agentic AI because autonomous systems may contain numerous interacting identities rather than a single application principal. Recent work specifically addressing an identity fabric for agentic AI emphasizes the need to design, implement, and evaluate identity security as an integrated architectural capability rather than as an isolated authentication service (Pappu, Bhushan, and Jaiswal, 2026).

A theoretical basis for adaptive identity management can also be derived from the supplied literature on

representation learning and adaptive computational systems. Neural-network research demonstrates how complex system behavior can be represented through multiple interacting features and hierarchical transformations (Jiao, Yang, Liu, Wang, and Feng, 2016). Similarly, feature-selection research demonstrates the importance of identifying informative characteristics dynamically rather than treating every input dimension as equally relevant (Liyanage, Zois, and Chelmis, 2021). These principles provide useful conceptual foundations for security architectures in which identity decisions depend on multiple dynamically changing attributes.

The central problem addressed in this study is therefore the absence of a unified conceptual mechanism through which agent identities can be continuously authenticated, evaluated, authorized, and adapted according to changing operational conditions. The objective is to develop an Adaptive Identity Fabric (AIF) that combines these functions while preserving clear separation between identity, trust, and authorization.

The paper has four principal objectives: first, to establish an architectural model for adaptive agent identity; second, to define a trust-management mechanism for continuous evaluation; third, to formulate a security-evaluation model for autonomous agent environments; and fourth, to analyze the benefits and limitations of adaptive identity management.

LITERATURE REVIEW

The supplied literature does not directly form a conventional body of agentic-AI identity research; instead, it provides complementary theoretical foundations involving representation, multiscale analysis, feature selection, neural computation, and adaptive classification. This distinction is important because it prevents unrelated technical findings from being incorrectly presented as direct evidence of identity-security performance.

Jiao's foundational works on neural-network system theory and neural-network computing establish a theoretical perspective in which complex computational behavior can be modeled through interconnected processing structures (Jiao, 1990; Jiao, 1993). For an identity fabric, this principle suggests that identity should not be represented as a single credential. Instead, an agent's security state can be modeled through multiple correlated attributes, including authenticated identity, workload context, behavioral history, requested operation, and environmental conditions.

The broader review of neural networks further supports the relevance of hierarchical computational representations for complex decision-making (Jiao, Yang, Liu, Wang, and Feng, 2016). In an adaptive identity fabric, hierarchical representation can conceptually separate low-level observations from higher-level security conclusions. For example, individual authentication events may constitute low-level evidence, while a continuously evaluated trust state represents a higher-level security abstraction.

The literature on multiscale and multiwavelet analysis provides another useful conceptual analogy. Multi-scale representation learning considers information at different levels of abstraction (Bai, Gong, Zhao, Lei, Yan, and Gao, 2021), while multiwavelet neural networks demonstrate how multiple representations can contribute to computational approximation (Jiao, Pan, and Fang, 2001). Applied to identity management, these principles support evaluating an agent at several levels: identity level, session level, action level, and ecosystem level.

Adaptive feature selection is particularly relevant to the proposed model. Liyanage, Zois, and Chelmis (2021) investigate dynamic instance-wise feature selection and classification, suggesting that relevant features can vary according to individual instances. An identity fabric can similarly adapt its decision inputs according to context. A low-risk read operation may require relatively limited contextual evidence, whereas an agent requesting privileged administrative access may require substantially richer evidence.

Other supplied studies demonstrate the value of extracting discriminative information from complex signals. Deep feature representation is used in imitation learning for autonomous helicopter aerobatics (Chen, Cao, Kang, Li, and Sun, 2021), while hierarchical feature prediction is investigated for fine-grained image recognition (Yu, Tan, Zhang, Tao, and Rui, 2019). These studies do not establish security mechanisms themselves, but they reinforce the broader theoretical proposition that system decisions can benefit from structured, hierarchical representations.

Wavelet and multiscale approaches further demonstrate the importance of transforming complex information into useful representations (Zhang, Zhou, and Jiao, 2004; Do and Vetterli, 2003). Within identity security, analogous transformations can convert raw telemetry into interpretable trust indicators.

The principal research gap is therefore architectural rather than algorithm-specific. Existing supplied research provides mechanisms for representation and adaptive processing, but does not establish a complete identity lifecycle spanning authentication, continuous trust assessment, authorization, delegation, and revocation for autonomous agents. The

identity-fabric perspective of Pappu, Bhushan, and Jaiswal (2026) directly motivates addressing this gap through an integrated security architecture.

METHODOLOGY

3.1 Adaptive Identity Fabric Architecture

The proposed AIF consists of six logically interconnected layers:

1. Identity Establishment Layer
2. Credential and Attestation Layer
3. Context and Telemetry Layer
4. Trust Evaluation Layer
5. Policy and Authorization Layer
6. Lifecycle and Security Governance Layer

The Identity Establishment Layer assigns every autonomous agent a distinct logical identity. The identity is associated with an agent instance, software workload, organizational role, or controlled service rather than merely with a human operator. This distinction is essential because two agents operated by the same user may have different capabilities and risk profiles.

The Credential and Attestation Layer validates whether an agent is operating from an authorized computational environment. Authentication is consequently treated as the first security decision rather than the final security decision.

The Context and Telemetry Layer collects information required for adaptive evaluation. Relevant contextual dimensions may include requested resource, operational environment, historical behavior, interaction pattern, session characteristics, and policy sensitivity. The conceptual motivation follows adaptive feature-selection principles in which different characteristics may have different relevance for individual decisions (Liyanage, Zois, and Chelmis, 2021).

The Trust Evaluation Layer converts these observations into a dynamic trust state. A conceptual trust function can be expressed as:

Ta(t)=wII+wCC+wBB+wHH−wRRT_a(t)=w_I I+w_C C+w_B B+w_H H-w_R R

where Ta(t)T_a(t) represents the trust score of agent aa at time tt; II represents identity confidence; CC represents contextual confidence; BB represents behavioral consistency; HH represents historical reliability; and RR represents detected risk. The weights wI,wC,wB,wH,wRw_I,w_C,w_B,w_H,w_R are policy-dependent.

The equation is not intended as a universal numerical standard. Rather, it demonstrates the central methodological principle: trust should be multidimensional and continuously recalculated. This approach is conceptually consistent with research emphasizing multiple feature representations and adaptive computational decision-making (Jiao, Yang, Liu, Wang, and Feng, 2016).

3.2 Risk-Adaptive Authorization

Authentication establishes identity, while authorization determines permissible actions. The proposed framework therefore prevents a valid identity from automatically receiving unrestricted authority.

A policy engine evaluates:

A=f(T,P,C,O)A = f(T, P, C, O)

where AA is the authorization decision, TT is current trust, PP is applicable policy, CC is contextual information, and OO is the requested operation.

For example, an agent with high trust may access a normal internal dataset but still be denied access to a highly sensitive repository if its current context violates policy. Conversely, a low-risk task can receive limited permissions even when the overall agent has broader organizational capabilities.

This separation reduces the impact of credential compromise because possession of a valid identity does not automatically imply unlimited authorization.

3.3 Continuous Identity Adaptation

An important methodological characteristic of AIF is continuous reassessment. Static identity systems commonly evaluate security at login or connection establishment. AIF instead reevaluates the agent when significant contextual events occur.

For instance, an agent performing document classification may initially receive read-only access. If it subsequently attempts to invoke a privileged administrative tool, the system can trigger additional identity verification and trust assessment. If behavioral evidence becomes inconsistent with the established operational profile, authorization can be reduced or suspended.

This adaptive process is conceptually related to dynamic feature selection, where the relevance of information can change according to the specific instance (Liyanage, Zois, and Chelmis, 2021).

3.4 Delegation and Agent-to-Agent Trust

Agentic ecosystems frequently involve delegation. Agent A may ask Agent B to execute a task, which creates a transitive security problem: should Agent B inherit Agent A's authority?

The proposed fabric does not automatically transfer complete privileges. Instead, delegated authority is represented as a constrained capability:

D={Is,Id,O,S,E,Tmin}D = \{I_s,I_d,O,S,E,T_{min}\}

where IsI_s represents the source agent, IdI_d the delegated agent, OO the permitted operation, SS the scope, EE the expiration condition, and TminT_{min} the minimum trust requirement.

This structure prevents unrestricted privilege propagation. If Agent B's trust level falls below TminT_{min}, the delegated authority becomes invalid.

3.5 Security Evaluation Model

The proposed evaluation assesses five dimensions: authentication resilience, authorization accuracy, trust responsiveness, containment capability, and operational overhead.

Hypothetically, an experimental environment could contain legitimate agents, compromised agents, malicious delegation attempts, abnormal tool invocations, and policy violations. Evaluation scenarios would compare static identity management against AIF.

Security effectiveness can be represented as:

SE=αD+βC+γR−δOSE = \alpha D + \beta C + \gamma R - \delta O

where DD represents detection effectiveness, CC containment effectiveness, RR response adaptability, and OO operational overhead.

The methodology deliberately evaluates both security benefits and costs. A system that detects every anomaly but imposes excessive computational or administrative overhead would not constitute an optimal enterprise identity architecture.

RESULTS

The conceptual evaluation indicates that the proposed Adaptive Identity Fabric provides five principal security improvements.

First, identity becomes continuously meaningful rather than temporally static. Traditional authentication can establish that an entity possessed valid credentials at a particular moment, whereas AIF evaluates whether the identity remains consistent with current context and behavior. This distinction is particularly important for autonomous agents whose behavior can change after initial authentication. The identity-fabric direction is therefore more appropriate for agentic environments in which autonomous execution extends beyond conventional user sessions (Pappu, Bhushan, and Jaiswal, 2026).

Second, adaptive trust reduces excessive dependence on individual credentials. If an agent's credential is stolen, a static model may continue granting access until the credential is revoked. AIF can detect inconsistencies through contextual and behavioral signals and progressively reduce authorization. This reflects the broader computational principle that meaningful decisions can depend on combinations of features rather than one isolated variable (Liyanage, Zois, and Chelmis, 2021).

Third, multilevel representation improves policy granularity. Identity can be considered simultaneously at workload, session, action, and ecosystem levels. Such hierarchical representation is conceptually compatible with the supplied research on hierarchical and multiscale feature processing (Bai, Gong, Zhao, Lei, Yan, and Gao, 2021; Yu, Tan, Zhang, Tao, and Rui, 2019). In practical terms, an agent may be trusted at the workload level but restricted for a particular sensitive action.

Fourth, delegation becomes controllable rather than implicitly transitive. The proposed constrained delegation model limits authority by operation, scope, duration, and minimum trust. This provides a mechanism for containing compromised or malfunctioning agents within multi-agent workflows.

Fifth, security evaluation becomes multidimensional. Rather than measuring only authentication success rates, the framework evaluates detection, containment, adaptation, and operational cost. This is important because security decisions in agentic AI involve dynamic system states rather than isolated login events.

However, the results also reveal limitations. Continuous evaluation requires additional telemetry, policy computation, and trust-management infrastructure.

Incorrect trust estimation can create either excessive restrictions or insufficient protection. Furthermore, adaptive systems may become difficult to govern when hundreds or thousands of agents interact simultaneously. The architecture therefore improves theoretical security resilience but does not eliminate the need for careful policy engineering, observability, calibration, and governance.

Overall, the findings support the proposition that identity security for autonomous AI should evolve from credential-centered authentication toward identity-plus-context-plus-trust decision-making.

DISCUSSION

The primary theoretical contribution of the proposed framework is the conceptual separation of identity, trust, and authorization. These concepts are frequently conflated in simpler security architectures. An authenticated agent is not necessarily trustworthy for every operation, and a trustworthy agent is not necessarily authorized for every resource. AIF explicitly preserves these distinctions and consequently provides a more precise foundation for agentic security.

The architecture also extends the significance of adaptive representation concepts from the supplied literature into identity management. Research on neural systems, feature representation, and multiscale processing demonstrates that complex computational decisions can benefit from structured representations rather than single-dimensional inputs (Jiao, 1990; Jiao, 1993; Jiao, Yang, Liu, Wang, and Feng, 2016). In the identity context, the equivalent principle is that security decisions should integrate multiple dimensions of evidence.

A significant practical implication is improved containment. Consider an enterprise agent that has legitimate access to customer-service data but unexpectedly attempts to invoke a financial administration function. A static authorization model might permit the request if the agent possesses a valid role. AIF can instead identify the mismatch between identity, context, historical behavior, and requested operation and initiate additional verification or deny the action. The framework thus moves security from permission ownership toward permission appropriateness.

Nevertheless, adaptivity introduces an important trade-off. Increasing the number of contextual factors can improve decision precision but also increases computational complexity and the possibility of false positives. Feature-selection research illustrates why dynamically identifying relevant information is valuable (Liyanage, Zois, and Chelmis, 2021), but an identity system must additionally ensure that the selected signals are reliable and resistant to manipulation.

Another challenge concerns trust calibration. Trust scores can create an appearance of mathematical precision even when underlying evidence is uncertain. Therefore, trust values should be interpreted as decision-support indicators rather than absolute measurements of agent morality or reliability. High trust should not permanently eliminate authorization controls.

The framework also raises governance questions. Autonomous agents may create new agents, delegate tasks, or operate across organizational boundaries. Consequently, identity lifecycle management must address creation, registration, credential rotation, suspension, revocation, delegation, and retirement. The identity-fabric approach is valuable precisely because these functions can be treated as one coordinated lifecycle rather than disconnected security controls (Pappu, Bhushan, and Jaiswal, 2026).

The supplied literature also reveals an important research limitation: most cited works concern computational representation rather than identity security. Their contribution to this paper is therefore theoretical and methodological rather than empirical evidence for AIF's security performance. The proposed framework should consequently be validated through controlled experiments involving real agentic workloads, attack simulations, latency measurements, false-positive rates, and authorization accuracy.

Future evaluation should compare alternative trust models, investigate adversarial manipulation of telemetry, and quantify the computational cost of continuous identity decisions. Such work would transform the current conceptual framework into a measurable enterprise security architecture.

CONCLUSION

This paper proposed an Adaptive Identity Fabric for Agentic AI Systems that integrates identity establishment, contextual assessment, continuous trust management, adaptive authorization, constrained delegation, and identity lifecycle governance. The central argument is that autonomous AI agents require security mechanisms capable of evaluating not only whether an identity is legitimate, but also whether its current action is appropriate within a particular context.

The framework uses a multidimensional trust model in which identity confidence, contextual evidence, behavioral consistency, historical reliability, and risk contribute to security decisions. It further separates authentication from

authorization and prevents unrestricted privilege propagation during agent-to-agent delegation.

Theoretical foundations were derived from the supplied literature on neural systems, feature representation, adaptive feature selection, hierarchical processing, and multiscale analysis. These works support the conceptual treatment of identity as a structured and adaptive representation rather than a single credential. The identity-fabric perspective provides the direct architectural motivation for applying these principles to agentic AI security (Pappu, Bhushan, and Jaiswal, 2026).

The proposed architecture offers potential advantages in credential-compromise containment, contextual authorization, trust adaptation, and multi-agent governance. However, its effectiveness depends on accurate telemetry, appropriate policy design, reliable trust calibration, and acceptable operational overhead. Future research should therefore implement the framework in realistic multi-agent environments and evaluate detection accuracy, authorization latency, false-positive rates, resilience against compromised agents, and scalability.

Ultimately, the evolution from static authentication toward an adaptive identity fabric represents a significant architectural direction for securing autonomous AI ecosystems. The principal contribution of AIF is not another isolated authentication mechanism, but a coordinated model in which identity, trust, context, authorization, and lifecycle management continuously interact to regulate autonomous behavior.

REFERENCES

1. J. Bai, B. Gong, Y. Zhao, F. Lei, C. Yan, and Y. Gao, “Multi-scale representation learning on hypergraph for 3D shape retrieval and recognition,” IEEE Trans. Image Process., vol. 30, pp. 5327–5338, May 2021.
2. Z. Cao, K. Zhang, and J. Wu, “FPB: Improving multi-scale feature representation inside convolutional layer via feature pyramid block,” in Proc. IEEE Int. Conf. Image Process., 2020, pp. 1666–1670.
3. S. Chen, Y. Cao, Y. Kang, P. Li, and B. Sun, “Deep feature representation based imitation learning for autonomous helicopter aerobatics,” IEEE Trans. Artif. Intell., vol. 2, no. 5, pp. 437–446, Oct. 2021.
4. M. N. Do and M. Vetterli, “The finite ridgelet transform for image representation,” IEEE Trans. Image Process., vol. 12, no. 1, pp. 16–28, Jan. 2003.
5. Y. Duan, F. Liu, L. Jiao, P. Zhao, and L. Zhang, “SAR image segmentation based on convolutional-wavelet neural network and Markov random field,” Pattern Recognit., vol. 64, pp. 255–267, 2017.
6. L. Jiao, B. Hou, W. Shuang, and L. Fang, Image Multiscale Geometric Analysis: Theory and Applications. Xi’an, China: Xian Electron. Sci. Technol. Univ. Press, 2008.
7. L. Jiao, J. Pan, and Y. Fang, “Multiwavelet neural network and its approximation properties,” IEEE Trans. Neural Netw., vol. 12, no. 5, pp. 1060–1066, Sep. 2001.
8. L. Jiao, Neural Network Computing. Xi’an, China: Xidian Univ. Press, 1993.
9. L. Jiao, Neural Network System Theory. Xi’an, China: Xidian Univ. Press, 1990.
10. L. Jiao, S. Yang, F. Liu, S. Wang, and Z. Feng, “Seventy years of neural networks: Review and prospect,” Chin. J. Comput., vol. 39, no. 8, pp. 1697–1716, 2016.
11. Y. W. Liyanage, D.-S. Zois, and C. Chelmis, “Dynamic instance-wise joint feature selection and classification,” IEEE Trans. Artif. Intell., vol. 2, no. 2, pp. 169–184, Apr. 2021.
12. D. Luengo, D. Meltzer, and T. Trigano, “Sparse ECG representation with a multi-scale dictionary derived from real-world signals,” in Proc. 41st Int. Conf. Telecommun. Signal Process., 2018, pp. 1–5.
13. J. Shi, Y. Zhao, W. Xiang, V. Monga, X. Liu, and R. Tao, “Deep scattering network with fractional wavelet transform,” IEEE Trans. Signal Process., vol. 69, pp. 4740–4757, Jul. 2021.
14. J. Yu, M. Tan, H. Zhang, D. Tao, and Y. Rui, “Hierarchical deep click feature prediction for fine-grained image recognition,” IEEE Trans. Pattern Anal. Mach. Intell., to be published, doi: 10.1109/TPAMI.2019.2932058.
15. L. Zhang, W. Zhou, and L. Jiao, “Wavelet support vector machine,” IEEE Trans. Syst., Man, Cybern., Part B. (Cybern.), vol. 34, no. 1, pp. 34–39, Feb. 2004.
16. D. Zhou, R. Duan, L. Zhao, and X. Chai, “Single image super-resolution reconstruction based on multi-scale feature mapping adversarial network,” Signal Process., vol. 166, 2020, Art. no. 107251.
17. K. Pappu, B. Bhushan and N. Jaiswal, "Future-Proofing Identity Security for Agentic AI Systems: Design, Implementation, and Evaluation of an Identity Fabric," 2026 International Conference on Artificial Intelligence, Systems, and Emerging Technologies (ICAISET), Cairo, Egypt, 2026, pp. 1-7, doi: 10.1109/ICAISET66439.2026.11541783.