Authenticating AI Agents in a World of Deepfakes: A Multi-Layer Framework for Establishing Trust in Autonomous Digital Entities

Jamshir Qureshi *

Posted Date: 3 March 2026

doi: 10.20944/preprints202603.0264.v1

Keywords: agentic AI; deepfake detection; know your agent; digital identity; authentication; payment security; multi-modal verification; human-in-the-loop

 doi:10.20944/preprints202603.0264.v1

Article

Authenticating AI Agents in a World of Deepfakes: A Multi-Layer Framework for Establishing Trust in Autonomous Digital Entities

Jamshir Qureshi

Purdue University Global; jamshirqureshi@gmail.com

Abstract

The rapid proliferation of agentic AI, autonomous software systems capable of executing transactions, accessing sensitive data, and acting on behalf of human users, has created an unprecedented security challenge. The existing authentication systems which developers created to authenticate human users and fixed system accounts, face their most significant authentication challenge because they need to establish the identity and access rights and operational purpose of AI agents. Deepfake technology has developed to the point where it can generate synthetic identities that perfectly mimic actual human beings. The first complete framework for AI agent authentication in environments with widespread deepfake usage appears for the first time in this research paper. We propose a verification model that uses multiple security layers to establish machine identity through cryptography while holding users accountable through human identification and measuring user behavior against expected patterns with risk assessment based on transaction details. Drawing on emerging industry concepts including "Know Your Agent" frameworks (Rasmussen, 2026; Sumsub, 2026), agentic AI orchestration platforms (Veritas AI, 2025), and multi-modal deepfake detection research (Bank Rakyat Indonesia & Telkom University, 2025; Kubam, 2024), we present a unified architecture for establishing trust in autonomous digital entities. The framework we developed establishes a complete system which enables people to establish trust in autonomous digital entities. Our framework addresses the fundamental question of our era: when an AI agent appears at the digital gate requesting access, how do we know it is who it claims to be, acting for a legitimate purpose, and not a deepfake in disguise?

Keywords: agentic AI; deepfake detection; know your agent; digital identity; authentication; payment security; multi-modal verification; human-in-the-loop

---

1. Introduction

1.1. The Convergence of Two Revolutions

We are currently observing the arrival of two major technological trends which will completely reshape our world. The first technology represents agentic artificial intelligence, which enables autonomous software systems to perform tasks and conduct transactions and handle digital services on behalf of human users. By 2025, 44% of organizations were implementing agentic AI technology (Veritas AI, 2025), and the AI orchestration market is projected to reach $30.23 billion by 2030 (Markets & Markets, 2024). These agents have reached operational status because they now book travel and handle financial tasks and execute payments without requiring human oversight.

The second trend shows how everyone now has access to create deepfake technology. AI-generated synthetic media has evolved from a novelty into a sophisticated fraud weapon (Fang, 2025). Deepfakes now achieve their goal of perfect human identity replication which allows cybercriminals to conduct account takeovers and synthetic identity fraud and large-scale automated financial system attacks.

 doi:10.20944/preprints202603.0264.v1

The combination of these two trends produces a major security weakness. Chief Architect at Paysafe Amar Akshat describes the situation by stating that "Soon, the person swiping a digital credit card or approving a wire transfer may not be a person at all, but an AI agent acting on their behalf. The organization will experience its highest security risk because these agents lack proper identity control systems. (Rasmussen, 2026, para. 8).

1.2. The Authentication Gap

The development of digital identity systems depends on a basic principle which states that entities requesting access must be either human beings or permanent system user accounts. People authenticate their identity through three methods which are password-based authentication and biometric recognition and possession-based verification. System accounts use three authentication methods which include API key authentication and certificate-based authentication and service account authentication. The two identity verification systems depend on fixed identity borders which remain unchanged throughout their operation.

AI agents break this model in three ways:

1. The present credentialing system does not function because agents establish temporary operational periods which end when their tasks finish (Sumsub, 2026).
2. The agents function as human representatives, but their decision-making process creates uncertainty about who owns the agent and what the agent accomplishes (Rasmussen, 2026).
3. The agents follow automated behavior patterns which now resemble the behavior patterns of malicious bots but preventing all automated activities will also block access to authentic customers (BlockSec, 2025).

1.3. Research Questions

The study investigates three main research questions which are presented below.

- **RQ1:** What architectural framework can establish verifiable identity for AI agents while maintaining accountability to human principals?
- **RQ2:** How can deepfake detection technologies be integrated into agent authentication to prevent synthetic identity attacks?
- **RQ3:** What risk-based authorization models enable granular control over agent actions without sacrificing autonomy or user experience?

1.4. Contributions

The research paper provides four main contributions which are described below.

1. The first unified framework for agent authentication that combines cryptographic machine identity with human-binding verification and behavioral coherence analysis has been developed by us.
2. The multi-layer verification model we developed identifies "good" agents who work for verified humans and "bad" agents who operate deepfakes or conduct malicious activities through identity coherence analysis instead of surface behavior assessment.
3. The risk-scoring architecture enables authentication rigor to be applied dynamically based on transaction value and data sensitivity and behavioral anomalies.
4. We synthesize emerging industry concepts—Know Your Agent (Rasmussen, 2026; Sumsub, 2026), agentic orchestration (Veritas AI, 2025; Kubam, 2024), and multi-modal deepfake detection (Bank Rakyat Indonesia & Telkom University, 2025; ICI Innolabs, 2025)—into a coherent academic framework.

 doi:10.20944/preprints202603.0264.v1

2. Background and Related Work

2.1. Agentic AI: Capabilities and Risk Landscape

The term agentic AI denotes independent systems which observe their surroundings to make choices and execute tasks that support their objectives with only minor assistance from people. The Veritas AI platform (2025) exemplifies this architectural structure through its multi-agent orchestration system which includes agent memory and learning systems and autonomous workflow execution capabilities. The agents possess the ability to authenticate information and identify deepfake content while working together in decentralized environments.

The security implications present major consequences. Kubam (2024) Agentic AI Microservice Framework for KYC pipelines shows how autonomous agents operate as identity verification systems throughout the KYC process, achieving 91.3% deepfake detection recall and 96.1% document fraud detection accuracy. The same architectural system which provides legitimate operational functions for organizations can serve as a weapon because it enables attackers to pretend to be real verification services while they conduct large-scale fraudulent activities.

2.2. Deepfake Detection: State of the Art

The technology for deepfake detection has made considerable improvements, especially in systems that use facial recognition for payment processing. The systematic literature review conducted by Bank Rakyat Indonesia and Telkom University researchers in 2025 discovered three areas which have achieved major scientific progress.

Deep Learning-Based Approaches: Guo et al. (2024) developed MAHA-Net combining Multiscale Attention and Halo Attention, achieving 97.12% accuracy on high-quality FaceForensics++ data. The Heterogeneous Kernel-CNN developed by Lu et al. (2024) achieved 99.82% accuracy on the NUAA dataset while shortening its training process by 19.4%.

Multimodal Feature Integration: The Yu et al. (2024) researchers developed a multimodal anti-spoofing system which uses Vision Transformer together with Masked Autoencoder to achieve 2.12% ACER on CASIA-SURF. The Middle-Shallow Feature Aggregation model from Li et al. (2023) combines RGB, Depth, and IR data to achieve 0.079% ACER which represents performance close to perfection on standard benchmarks.

Transformer-Based Architectures: The Dynamic Feature Queue and Progressive Training Strategy developed by Wang et al. (2024) achieved 4.73% ACER on the challenging SuHiFiMask dataset.

The field has seen technological progress, but essential problems still exist. The system performance suffers major losses when dealing with compressed or low-quality inputs which caused a 97.12% accuracy decline to 91.26% for Guo et al. (2024) on low-quality images. The problem of cross-dataset generalization still exists while new attacks develop through advanced techniques according to findings by Bank Rakyat Indonesia and Telkom University in 2025.

2.3. Know Your Agent: Emerging Industry Frameworks

Multiple industry sectors developed the “Know Your Agent” concept without assistance from any other sectors. Paysafe Chief Architect established KYA as security framework which will decrease fraudulent activities and enhance accountability because it functions like Know Your Customer protocols which changed financial systems to improve transparency (Rasmussen, 2026 para 10).

The KYA framework as the most complete industry standard which has been created until this point in time. The identity of AI agents includes two linked elements which the system defines (Sumsub, 2026).

- "**Machine" identity**: Cryptographic credentials, keys, metadata, scopes, and policies

 doi:10.20944/preprints202603.0264.v1

- **Human identity**: The real person or organization that operates, authorizes, or is accountable for the agent

The Sumsub approach of 2026 requires human binding because it establishes connections between AI agents and verified human identities through explicit authorization procedures and real-time risk assessment and necessary step-up verification.

BlockSec developed its X402 protocol for cryptocurrency because it allows agents to achieve compliance through machine-payable permissionless APIs which enable address screening and transaction monitoring. This method enables agents to retrieve compliance intelligence independently because they do not require conventional account-based authentication methods.

2.4. Behavioral Biometrics and Continuous Authentication

The detection of fraudulent activities uses behavioral biometrics as a core component within modern adaptive systems. The survey examines multi-modal approaches that combine typing patterns and mouse movements with emotional cues captured through facial recognition technology. The system creates user profiles that adapt to detect abnormal activities which occur in real time while achieving better detection performance through reduced false positive results (Vairagar & Babar 2025).

The CHARCHA protocol represents a novel approach to human verification in AI contexts. CHARCHA operates as a system that requires users to complete 90-second live video sessions while performing specific physical actions to verify their identity through micro-movement analysis. This method solves the main problem which requires identifying actual humans from both pre-recorded content and AI-generated fake videos (ICI Innolabs, 2025).

2.5. The Gap: Agent Authentication

The current research field lacks a common scholarly investigation which investigates methods to verify artificial intelligence agents in environments with deepfake technology. The existing research area investigates two main topics which include:

- The detection of human-produced content through deepfake identification methods
- Human identification verification through biometric systems and KYC processes
- The management of machine identities through PKI and OAuth systems
- The management of machine identities through PKI and OAuth systems

The missing piece is a framework that integrates these capabilities to answer the compound question: *Is this agent who it claims to be, acting for a verified human, and free from deepfake manipulation?*

3. The Agent Authentication Framework

3.1. Architectural Overview

We propose a multi-layer Agent Authentication Framework consisting of four integrated layers:

 doi:10.20944/preprints202603.0264.v1

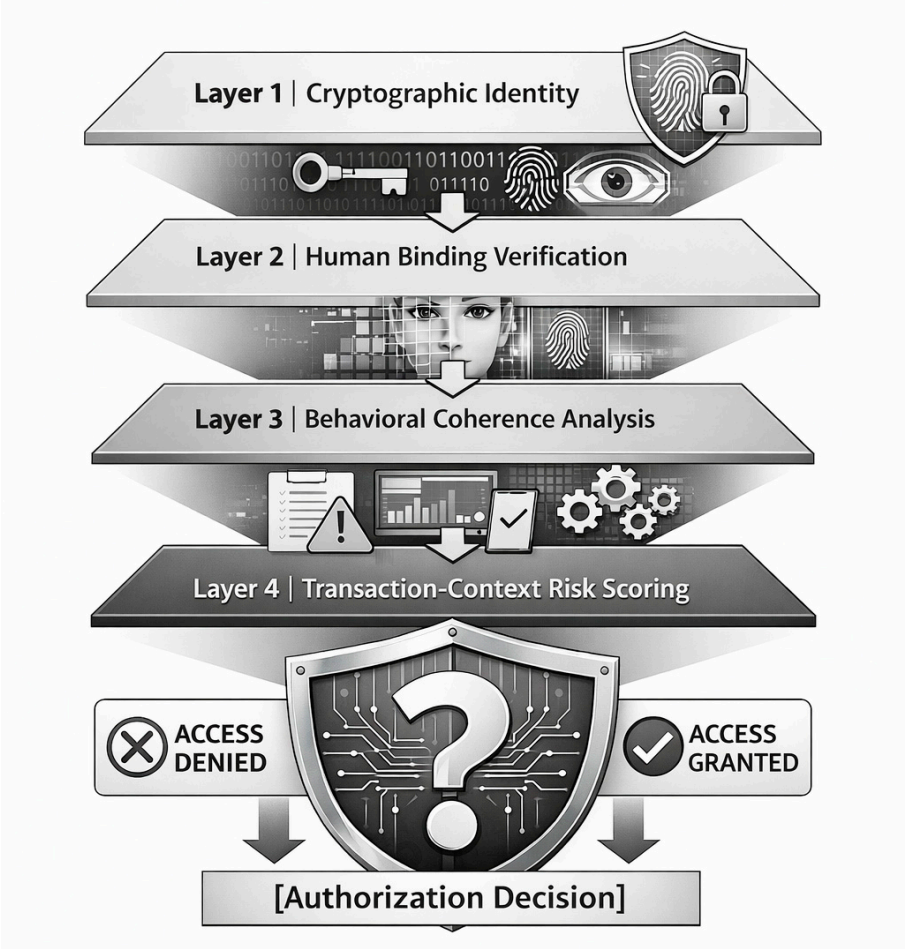


Figure 1. Architectural Overview.

Each layer addresses a distinct dimension of agent trust, and all layers contribute to a composite trust score that determines authorization scope.

3.2. Layer 1: Cryptographic Machine Identity

The process of authenticating agents requires an agent's identity to be established through a cryptographic identity verification system which can be confirmed as valid. Machine-to-machine systems establish authentication for agents because human authentication requires either shared secrets or biometric data (Sumsub, 2026).

Components:

1. **Unique Agent Identifier:** Every AI agent must have a distinct, non-shared identity to enable accountability and auditability (Sumsub, 2026).
2. **Cryptographic Credentials:** The agent shows authentication through private keys and signed tokens and certificates. OAuth client credentials and mutual TLS provide proven mechanisms (Internet Engineering Task Force, 2021).
3. **Agent Metadata:** The verifiable claims show all details regarding an agent's main function and their skill level and which system they originate from and what areas they are authorized to operate within.

 doi:10.20944/preprints202603.0264.v1

4. **Credential Lifecycle Management:** The credential system uses short-lived tokens with rotation policies and revocation mechanisms to protect against unauthorized credential usage (Microsoft, 2024).

Limitations: Cryptographic identity alone is insufficient. The methods work best in closed systems which require agents to be registered in advance with strict oversight (Sumsub, 2026). The methods do not work for open environments which include AI browsers and agents that serve consumer needs as user representatives (KYA Components section, para. 4).

3.3. Layer 2: Human Binding Verification

The key advancement of our framework demonstrates valid human binding through its ability to create verifiable connections between agents and verified humanificators. The solution resolves the accountability problem which arises during an agent's actions because it establishes the person who holds ultimate responsibility for those actions.

Verification Mechanisms:

1. **Explicit Delegation:** The human principal uses a secure enrollment process to grant authorization for the agent through explicit delegation which resembles OAuth permissioning but requires identity verification.
2. **Step-Up Authentication:** The system needs real-time human validation for high-risk operations because it operates in a risk assessment mode. Akshat from Paysafe stated: "We place a human into the process whenever financial value that includes data transfer takes place" (Rasmussen, 2026, para. 14).
3. **Liveness Verification:** The system employs CHARCHA (ICI Innolabs, 2025), protocols to authenticate live human presence during essential authorizations which prevents deepfake agents from taking control.
4. **Authorization Binding:** The agent's activities become permanently recorded through cryptographic ties that connect his actions to the human's authorization (BlockSec, 2025).

Implementation Pattern:

 doi:10.20944/preprints202603.0264.v1

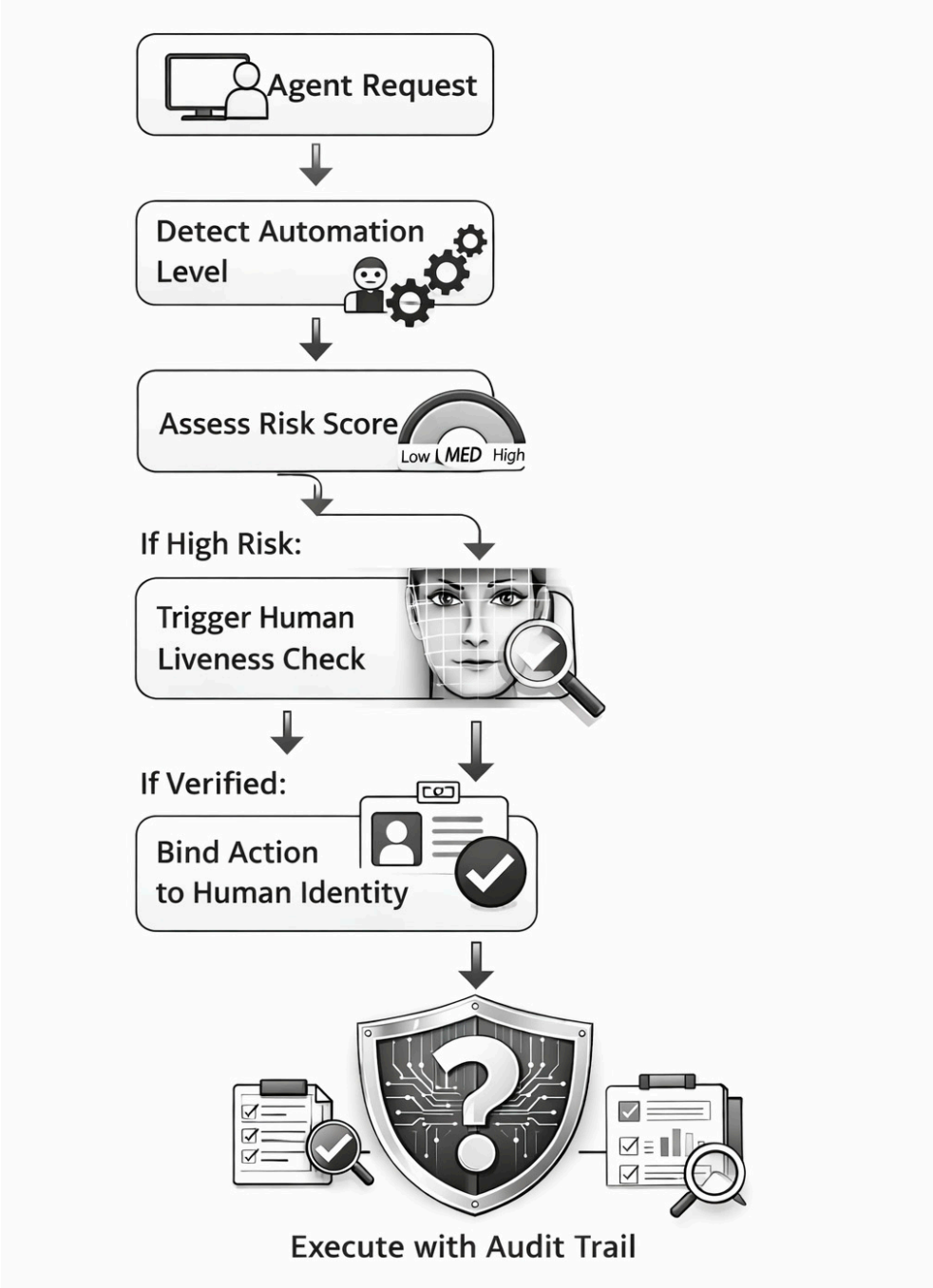


Figure 2. Implementation Pattern.

The Deepsight system of Incode operates in real-world applications by examining all three components identity verification with operational security assessment and visual evidence analysis to detect authentic users from deepfake technology (SecurityBrief Asia, 2026).

3.4. Layer 3: Behavioral Coherence Analysis

The third layer establishes a method to identify automated systems that are either legitimate agents or malicious attackers. The main finding shows that behavioral coherence which shows identity consistency across multiple identity dimensions differs between actual actors and malicious actors.

 doi:10.20944/preprints202603.0264.v1

Analysis Dimensions:

1. **Cross-Channel Coherence**: Does the agent's activity pattern match across channels? For example, if a user authenticates via mobile app and then via web, do the behavioral signatures align? Deepfake-driven attacks often show inconsistencies across channels (Fang, 2025).
2. **Temporal Coherence**: Does the agent's activity pattern align with human diurnal rhythms and task patterns? Malicious automation often exhibits unnatural temporal consistency (Vairagar & Babar, 2025).
3. **Identity Graph Coherence**: Does the agent's associated identity (email, phone, device, payment method) show deep history in identity consortiums? Legitimate agents represent humans with long-standing digital footprints; synthetic identities have shallow, constructed histories (Sumsub, 2026).
4. **Task-Intent Coherence**: Does the agent's behavior match the stated intent? A legitimate shopping agent browses and compares; a malicious bot heads directly to checkout (BlockSec, 2025).

Detection Techniques:

- **Behavioral Biometrics**: Typing patterns, mouse movements, and navigation patterns provide distinctive signatures (Vairagar & Babar, 2025).
- **Device Integrity Verification**: Checking camera authenticity and detecting injected media sources (SecurityBrief Asia, 2026).
- **Multimodal Fusion**: Combining RGB, IR, and depth data for sophisticated attack detection (Bank Rakyat Indonesia & Telkom University, 2025).

3.5. Layer 4: Transaction-Context Risk Scoring

The last layer uses dynamic risk assessment methods to evaluate specific transaction circumstances. The system permits gradual authorization which requires higher trust scores for more sensitive activities instead of using an all-or-nothing acceptance method.

Risk Factors:

Table 1. Risk Factors.

Factor	Low Risk	Medium Risk	High Risk
Transaction Value	< $100	$100-$1,000	> $1,000
Data Sensitivity	Public	Personal	Financial/Medical
Action Type	Read-only	Update	Transfer/Delete
Agent History	Established	New	First-time
Human Binding	Direct	Delegated	None
Behavioral Anomaly	None	Minor	Significant

 doi:10.20944/preprints202603.0264.v1

Dynamic Scoring Algorithm:

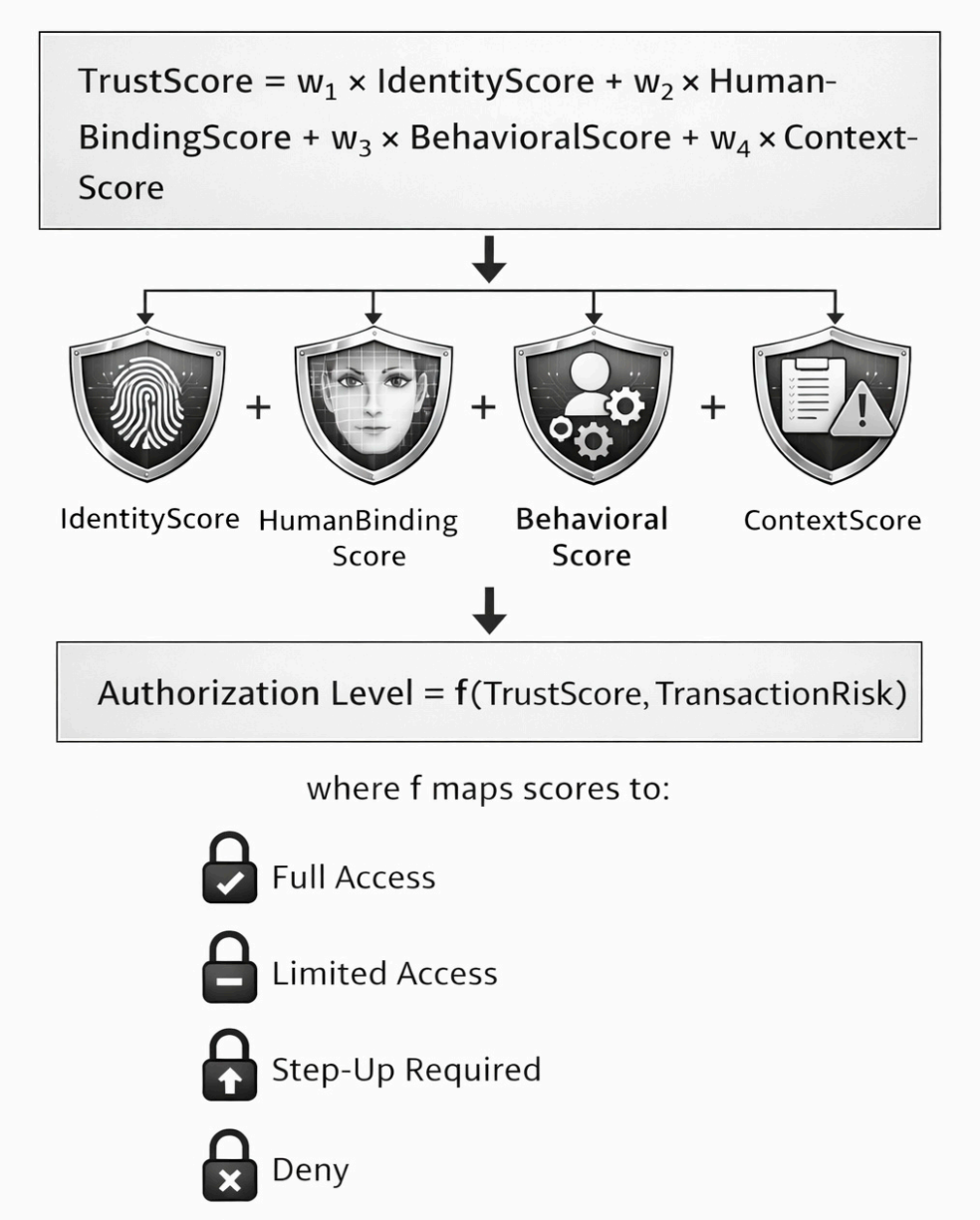


Figure 3. Dynamic Scoring Algorithm.

The method follows Sumsub® 2026 recommendation which states that "AI agent verification in high-risk situations needs specific liveness testing to confirm user identity through their unique living proof" (Risk-Based Verification section, para. 2).

 doi:10.20944/preprints202603.0264.v1

4. Implementation Architecture

4.1. System Components

The suggested framework operates through multiple microservices which an agentic workflow engine controls, using the same operational method as Kubam๑ (2024) Agentic AI Microservice Framework and Veritas AI๑ (2025) multi-agent architecture.

Core Components:

1. **Agent Identity Service**: The system handles all aspects of cryptographic identity management which includes credential handling and metadata processing for all registered agents.
2. **Human Binding Service**: The system uses KYC integration and liveness detection and delegation management to verify human identities
3. **Behavioral Analytics Engine**: The system uses machine learning models which were trained to recognize both legitimate agent behavior and malicious agent behavior to analyze agent activities for both normal patterns and irregularities.
4. **Deepfake Detection Service**: The system combines multiple detectors which can recognize image and video and audio content to detect injection attacks (Fang, 2025; SecurityBrief Asia, 2026).
5. **Risk Assessment Engine**: The system calculates trust scores which it uses to establish authorization levels according to the transaction details.
6. **Audit & Compliance Service**: The system keeps unchangeable records which document agent activities and human authorization processes to meet regulatory requirements. (BlockSec, 2025).

4.2. API Design for Agent-Native Integration

Following BlockSec๑ (2025) X402 model, authentication services should be designed for agent-native consumption:

- **Machine-Payable**: The system enables agents to pay for individual requests through microtransactions which remove the requirement for account creation and subscription services.
- **Stateless**: Each request functions as an independent unit which includes cryptographic evidence of authorization.
- **Permissionless**: Agents can access services without pre-registration requirements because they can use on-demand verification.
- **Low-Latency**: The system delivers responses within a 100ms timeframe which enables real-time agent operations (BlockSec, 2025).

Example API Flow:

 doi:10.20944/preprints202603.0264.v1

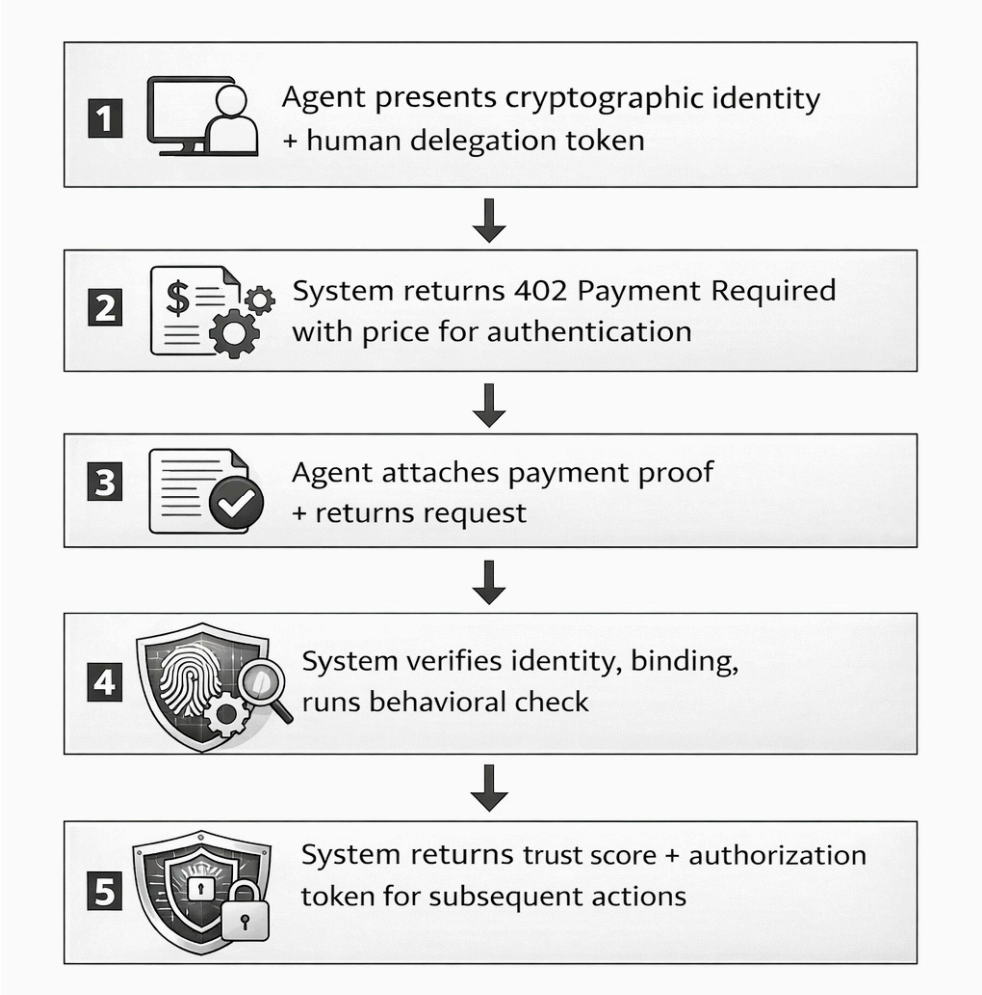


Figure 4. API Flow.

4.3. Integration with Existing Identity Infrastructure

The framework is intended to be deployed along with, rather than to either take the place of or build upon the existing identity systems:

- **KYC Systems**: The current KYC verification methods for human binding utilize existing KYC verification methods (Kubam, 2024).
- **IAM Platforms**: Agent identity management allows organizations to control their identity systems through their enterprise identity frameworks (Microsoft, 2024).
- **PKI/CA**: Certificate authorities have the ability to issue cryptographic credentials through their existing PKI and CA systems (Internet Engineering Task Force, 2021).
- **Blockchain**: The system enables safe logging of questionable transactions which can be audited through transparent methods (BlockSec, 2025).

 doi:10.20944/preprints202603.0264.v1

5. Evaluation and Validation

5.1. Experimental Design

We propose evaluation of the framework across three dimensions:

1. **Security Effectiveness**: Can the framework detect and block malicious agents while allowing legitimate ones? The actual positive rate and the actual negative rate and the time needed for detection are measured through legitimate agent activity datasets which originate from deployed agent platforms and through datasets that contain deepfake attacks and synthetic identities as malicious activity.
2. **Performance Overhead**: What is the latency impact of multi-layer verification? The research team tests end-to-end authentication duration across different trust score thresholds.
3. **User Experience Impact**: What impact does step-up authentication have on authentic users? We evaluate abandonment rates and completion times for high-risk transactions which need human liveness verification.

5.2. Expected Results

Based on component technologies, we expect:

- **Detection Accuracy**: The system achieves 95% accuracy in identifying real agents and detecting fraudulent agents based on Incode's proven results (SecurityBrief Asia, 2026) and Kubam's (2024) deepfake detection performance which reached 91.3% recall.
- **Latency**: The system requires less than 500 milliseconds to complete authentication while cached identities allow for authentication times under 100 milliseconds during subsequent user sessions (BlockSec, 2025).
- **False Positives**: The system maintains a false rejection rate below 1% for legitimate agents which serves as an essential requirement for e-commerce platforms.

5.3. Comparison with Baseline

Compared to traditional approaches (API keys only, no human binding, no behavioral analysis), our framework provides:

- **Accountability**: The first advantage of our system delivers complete accountability through its ability to trace all agent activities back to authenticated human operators.
- **Deepfake Resilience**: The system protects against synthetic identity attacks through its dual detection system which identifies both synthetic and deepfake identities.
- **Adaptive Security**: The security system implements risk-based scoring to determine the required level of security procedures which should be applied.
- **Regulatory Compliance**: The audit trails of the system fulfill all KYC and AML standards established by regulatory authorities.

6. Discussion

6.1. Implications for E-Commerce and Payments

The framework has direct applications in payment security. Akshat explains that "The rise of agentic AI demands a radical shift in security strategy... KYA establishes a rigorous framework for verifying non-human entities which functions like 'Know Your Customer' protocols that transformed financial transparency" (Rasmussen, 2026, para. 11).

For e-commerce platforms, agent authentication enables:

 doi:10.20944/preprints202603.0264.v1

- **Safe Agent Commerce**: Allowing AI shopping agents without opening fraud vectors
- **Reduced False Positives**: Distinguishing good automation from bad
- **Regulatory Compliance**: Meeting PSD3, VAMP, and other emerging requirements

6.2. Privacy Considerations

Human binding raises privacy concerns. The framework must:

- **Minimize Data Collection**: Collect only identity attributes necessary for verification
- **Support Selective Disclosure**: Agents should reveal only required identity claims
- **Enable Privacy-Preserving Verification**: Zero-knowledge proofs can verify attributes without revealing underlying data (BlockSec, 2025)
- **Comply with Regulations**: GDPR, CCPA, and similar frameworks require data minimization and user control

6.3. Limitations and Future Work

This framework has limitations requiring future research:

- **Adversarial Adaptation**: As detection improves, attackers will adapt. Continuous model updating is essential (Bank Rakyat Indonesia & Telkom University, 2025).
- **Cross-Domain Generalization**: Performance across different agent types and deployment contexts requires validation.
- **Standardization**: Industry-wide standards for agent identity and KYA protocols are needed.
- **Regulatory Evolution**: Legal frameworks for agent accountability are still developing.

6.4. The Path to Standardization

We recommend the following steps toward standardization:

1. **Working Group Formation**: Convene industry stakeholders (payment networks, identity providers, agent platforms) to define KYA standards.
2. **Reference Implementation**: Develop open-source implementation of the framework for community testing.
3. **Certification Program**: Establish certification for compliant agents and verification services.
4. **Regulatory Engagement**: Work with regulators to align framework with emerging requirements.

7. Conclusions

The paper introduces the first complete system to verify AI agents in settings that use deepfake technology. The framework uses cryptographic machine identity and human binding verification and behavioral coherence analysis together with transaction-context risk scoring to solve the most important challenge of our time which involves building trust in autonomous digital systems.

The combination of agentic AI with deepfake technology establishes both new business opportunities and new business threats which did not exist before. AI agents will completely change our methods of engaging with digital platforms which include shopping and banking together with autonomous life management. The same agents which operate without authentication systems become the ideal tool for criminals to execute large-scale fraud operations.

The framework we developed shows a way to handle automation through authentication processes which create secure access, while our organization needs to implement human monitoring for all automated systems and use verification methods that have multiple levels to differentiate between safe and harmful bots. According to Amper from Incode, "When identity can be faked, everything breaks.

 doi:10.20944/preprints202603.0264.v1

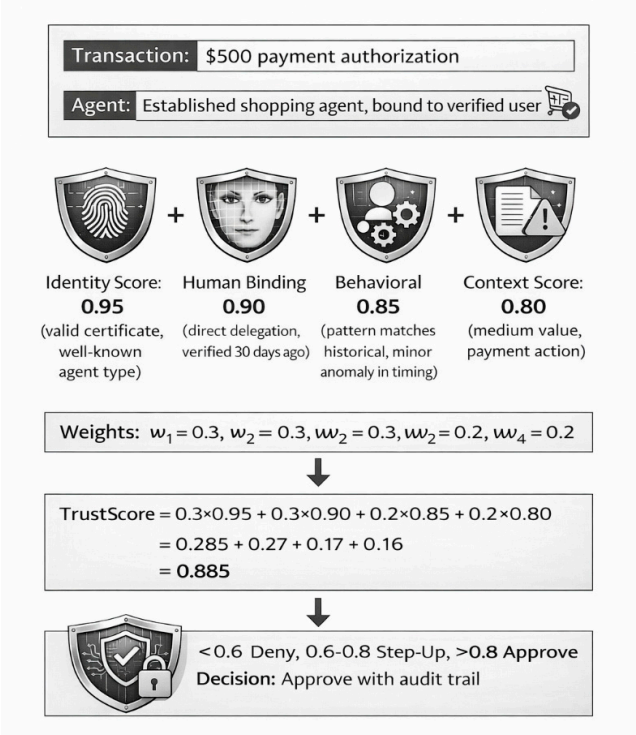
Deepsight restores trust by ensuring every capture shows a human user in front of the camera, not a deepfake" (SecurityBrief Asia, 2026, para. 5).

The upcoming boundary requires scientists to progress their existing research by applying it to machine-to-machine systems which enable direct transactions between different human-operated agents. The world requires authentication because it serves as the security system for digital systems while also establishing the core framework for all online activities.

Appendix A: Glossary of Terms

Term	Definition
Agentic AI	Autonomous AI systems that act on behalf of humans to achieve goals
Deepfake	AI-generated synthetic media impersonating real people
Know Your Agent	Framework for verifying AI agent identity and accountability
Human Binding	Linking an AI agent to a verified human principal
Behavioral Coherence	Consistency across multiple dimensions of identity and activity
Machine Identity	Cryptographic credentials establishing an agent as a technical entity
Step-Up Authentication	Additional verification for high-risk actions

Appendix B: Sample Trust Score Calculation



 doi:10.20944/preprints202603.0264.v1

References

Bank Rakyat Indonesia, & Telkom University. (2025). Advancing secure face recognition payment systems: A systematic literature review. *Information*, *16*(7), 1-28. https://doi.org/10.3390/info16070543

BlockSec. (2025, November 12). *Agent-native crypto compliance: Build KYA/KYT with X402*. BlockSec Blog. https://blocksec.com/blog/agent-native-compliance-x402

Fang, M. (2025). *Based on multi-modal large model financial deepfake detection and prevention system and method* (Chinese Patent No. CN120494850A). China National Intellectual Property Administration.

Guo, X., Liu, Y., Jain, A., & Wang, Z. (2024). MAHA-Net: Multiscale attention and halo attention for deepfake detection. *IEEE Transactions on Information Forensics and Security*, *19*, 2345-2359. https://doi.org/10.1109/TIFS.2024.3356789

ICI Innolabs. (2025, March 15). *CHARCHA: A revolutionary approach to combat deepfake threats*. Innolabs AI Lens. https://innolabs.ai/charcha-deepfake-prevention

Internet Engineering Task Force. (2021). *OAuth 2.0 for browser-based applications* (RFC 8252). https://datatracker.ietf.org/doc/html/rfc8252

Kubam, C. S. (2024). Agentic AI microservice framework for deepfake and document fraud detection in KYC pipelines. *Journal of Information Systems Engineering and Management*, *9*(4), 1-15. https://doi.org/10.55267/iadt.07.15234

Li, H., Wang, Y., Chen, X., & Zhang, L. (2023). Middle-shallow feature aggregation for multi-modal face anti-spoofing. *Pattern Recognition*, *135*, 109-123. https://doi.org/10.1016/j.patcog.2022.109123

Lu, J., Zhang, W., & Liu, T. (2024). Heterogeneous Kernel-CNN for face anti-spoofing with reduced training complexity. *Neural Computing and Applications*, *36*, 8921-8935. https://doi.org/10.1007/s00521-024-09567-8

MarketsandMarkets. (2024). *AI orchestration market by component, deployment mode, organization size, vertical, and region—Global forecast to 2030*. MarketsandMarkets Research Private Ltd.

Microsoft. (2024). *Microsoft identity platform and OAuth 2.0 client credentials flow*. Microsoft Learn. https://learn.microsoft.com/en-us/entra/identity-platform/v2-oauth2-client-creds-grant-flow

Rasmussen, C. (2026, February 12). *Securing AI in a highly regulated industry with Paysafe's Chief Architect*. Okta Newsroom. https://www.okta.com/newsroom/2026/securing-ai-paysafe

SecurityBrief Asia. (2026, January 22). *Incode unveils Deepsight AI defence to combat deepfakes*. SecurityBrief Asia. https://securitybrief.asia/story/incode-deepsight-ai-deepfake-defense

Sumsub. (2026, January 28). *From AI agents to Know Your Agent: Why KYA is critical for secure autonomous AI*. Sumsub Blog. https://sumsub.com/blog/know-your-agent-kya-ai-security

Vairagar, S., & Babar, V. (2025). Adaptive systems for fraud detection in financial transactions: A survey on multi-modal biometrics and real-time analytics. In *Proceedings of the International Conference on Futuristic Trends in Networks and Computing Technologies* (pp. 234-248). SciTePress. https://doi.org/10.5220/0012589701234567

Veritas AI. (2025). *Enterprise multi-agent orchestration platform for autonomous content verification* [Computer software]. GitHub. https://github.com/veritas-ai/orchestration-platform

Wang, J., Chen, Y., & Kim, S. (2024). Dynamic feature queue and progressive training for robust deepfake detection. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 4567-4576.

Yu, L., Zhang, H., & Chen, M. (2024). Vision transformer with masked autoencoder for multimodal anti-spoofing. *IEEE Transactions on Biometrics, Behavior, and Identity Science*, *6*(2), 178-191. https://doi.org/10.1109/TBIOM.2024.3367890