International Journal of Engineering Research and Science & Technology

www.ijerst.org

ISSN : 2319-5991

Vol. 22 No. 3 (2026)

ijerst.editor@gmail.com

editor@ijerst.com

ISSN 2319-5991 www.ijerst.com

Peer Reviewed, Referred & Indexed Journal

Research Paper

A SECURE DECENTRALIZED TRUST FRAMEWORK FOR AUTONOMOUS AI AGENTS USING BLOCKCHAIN, SMART CONTRACTS, AND ADAPTIVE REPUTATION LEARNING

Dr Prakash Krishna Shinde
Head of Computer Engineering Department, Dr D Y Patil Polytechnic Kolhapur, Maharastra

Abstract: Autonomous Artificial Intelligence (AI) agents have become a revolutionary force in distributed computing, enabling intelligent systems to independently perceive environments, make decisions, negotiate services, and collaborate without constant human intervention. Autonomous AI agents are increasingly being deployed in healthcare, finance, industrial automation, intelligent transportation, smart cities, decentralized finance (DeFi), and Internet of Things (IoT) ecosystems where secure and trustworthy interactions are of paramount importance. However, the existing trust management mechanisms mainly depend on centralized architectures that suffer from single points of failures, limited scalability, privacy concerns, and vulnerability to insider attacks. Blockchain technology provides a decentralized, immutable and transparent recording of transactions. But most of the blockchain-based trust management systems are constructed based on static reputation mechanisms which are not sufficient to describe the behavioral changes of autonomous AI agents in the course of time. Moreover, current frameworks rarely integrate adaptive trust learning and smart contract based policy enforcement which restricts their applicability in dynamic decentralized environments. We propose TrustChain-AI, a secure decentralized trust platform based on blockchain, smart contracts and adaptive reputation learning, for trustworthy interactions between autonomous AI agents. This paper proposes a framework which employs a permissioned blockchain to store immutable trust records and smart contracts to automate the processes of identity verification, transaction validation, policy enforcement, and trust updates. A dynamic reputation engine continually observes agent behavior based on interaction history, service reliability, peer feedback and behavioral consistency, and dynamically computes trust scores to detect malicious agents. Different from the traditional methods, TrustChain-AI combines the decentralized trust verification and the intelligent behavior assessment. The real-time trust adapting is allowed without the centralized supervision. Its aim is to give a framework for increased transparency, scaleability and resilience against identity spoofing, Sybil attacks, replay attacks and reputation manipulation. We present a reproducible experimental setup using a permissioned blockchain network and simulated autonomous AI agents to evaluate the prediction accuracy of trust, transaction throughput, consensus latency, detection of malicious agents and computational overhead. The comparative analysis demonstrates the effectiveness of the proposed framework in providing secure, scalable and adaptive trust management for next generation decentralized AI ecosystems.

Keywords: Autonomous AI Agents, Blockchain, Smart Contracts, Trust Management, Adaptive Reputation Learning, Decentralized Security, Distributed Ledger Technology, Multi-Agent Systems.

I. Introduction

AI has evolved from rule-based automations to intelligent systems that are capable of independent reasoning, adaptive learning, and collaborative decision-making. Recent advances in deep learning, reinforcement learning, large language models and multi-agent systems have boosted the development of autonomous AI agents that can autonomously perform complex tasks in continuous interaction with dynamic environments. Unlike traditional software applications that execute predefined sets of instructions, autonomous AI agents are able to sense changes in the environment, analyze contextual information, negotiate with peer entities, make strategic decisions, and learn to improve their performance via continuous learning. These capabilities have made autonomous AI agents an integral part of contemporary intelligent infrastructures including smart manufacturing, healthcare, autonomous transportation, cloud-edge computing, financial technology, cybersecurity, industrial Internet of Things (IIoT), and the like.

There are emerging decentralized intelligent systems that are changing the way computational entities cooperate and share information. In Industry 5.0, autonomous AI agents orchestrate robotic collaboration, production scheduling, predictive maintenance and quality assurance. Smart transportation systems' intelligent agents share traffic information, optimize route planning and collaborative autonomous driving. Financial institutions use AI agents for algorithmic trading, fraud detection and risk assessment. Healthcare organizations increasingly use intelligent diagnostic systems and clinical decision-support agents to improve patient outcomes. Autonomous service agents in cloud-edge computing dynamically allocate computational resources and optimize workload distribution across heterogeneous infrastructures. These applications demonstrate that autonomous AI agents are emerging as trusted digital actors executing safety-critical and financially sensitive operations.

But even with such technological advances, trust between autonomous AI agents remains a big problem. In distributed AI environments, interactions among independent agents from different organizations, administrative domains and geographical areas are common. Since these agents generally operate without central control, each interaction introduces an element of uncertainty regarding the authenticity, reliability, competence and behavioral integrity of the entities concerned. Traditional authentication approaches guarantee only the digital identity of an agent at the time of registration and are not able to guarantee trustworthy behavior during the operational lifetime of an agent. Thus, compromised authenticated agents may still take part in collaborative processes while showing malicious or unreliable behavior due to malware, adversarial manipulation, insider threats, or software vulnerabilities. These limitations severely limit the security and reliability of decentralized AI ecosystem.

Traditional trust management systems rely on a central authority that maintains reputation databases, authenticates users, authorizes transactions and implements security policies. Although the centralized architectures facilitate administration, they involve disadvantages like single points of failure, poor scalability, performance bottlenecks and lack of transparency. Centralized trust repositories are also attractive targets for cyber-attacks, because if the central authority is compromised, the security of the whole distributed environment can be compromised. As decentralized AI ecosystems grow, centralized trust management becomes less and less suitable to support millions of autonomous agents performing continuous interactions over globally distributed networks.

One of the most promising technologies to overcome these limitations is the blockchain technology which decentralizes the trust establishment through distributed ledger technology. Blockchain is a secure and immutable ledger of all transactions in the network, protected by cryptography and without the need for centralized intermediaries, thus providing transparency, integrity and accountability. Consensus protocols validate transactions and commit them on a distributed ledger replicated on a set of participating nodes. Blockchain's decentralized structure means data is less likely to be tampered with, unauthorized changes made, single-point failures, and resistance to cyberattacks improved.

Besides immutability of transactions, blockchains offer strong cryptographic identity management using public-key infrastructures and digital signatures. Each autonomous AI agent can have a unique identity on the blockchain, enabling secure authentication and non-repudiation of communication between agents. Smart contracts are also a tool to improve the function of blockchains, as they allow for programmable, self-executing contracts that automatically implement pre-defined security policies, without any human intervention. These contracts can be used to verify identities, access resources, refresh trust scores, capture evidence of conduct and execute collaborative agreements with full transparency. Hence, blockchain and smart contracts offer an appealing foundation for secure decentralized trust management.

But the blockchain doesn't fully address the trust problem. Current blockchain-based trust systems mostly emphasize on the secure recording of transactions and identity verification, but less on the ongoing behavioral evaluation of autonomous AI agents. Most of the existing systems assign static or slowly changing reputation values based on past interactions, which is not enough to detect fast evolving malicious behavior. Sophisticated adversarial agents can gradually distort trust values through coordinated attacks, collusive recommendations, or sporadic malicious activities that evade traditional reputation mechanisms. Therefore, future decentralized AI ecosystems require trust models that can continuously learn from behavioral evidence and dynamically adjust trust values to changes in interaction patterns.

Major Contributions

The principal contributions of this work are summarized below.

- Decentralized Trust Management Architecture Based on Permissioned Blockchain for Autonomous AI Agents.
- Adaptive reputation learning scheme to dynamically update trust values based on: behavioral consistency, transaction success rate, peer feedback and service reliability.
- Smart Contracts for automating identity verification, trust validation, policy enforcement and decentralized transaction authorization.
- Behavioral monitoring to identify malicious agents and automatically quarantine suspicious participants.
- A complete implementation methodology with reproducible experimental evaluation of simulated autonomous AI agents deployed on a permissioned blockchain network.
- Performance evaluation metrics such as trust prediction accuracy, transaction throughput, consensus latency, malicious agent detection rate, communication overhead and scalability
- Comparison analysis with enhanced existing blockchain-based trust management approaches.

2. Literature Survey

The rapid development of autonomous Artificial Intelligence (AI) agents has spurred the growth of decentralized intelligent systems that can independently reason, collaborate in decision-making, distribute resources and intelligently provide services. These systems are increasingly becoming part of heterogeneous environments with multiple AI agents interacting with each other in absence of central supervision, which raises major challenges in trust establishment, authentication, transparency, accountability and malicious behavior detection. With its distributed ledger architecture, cryptographic security, consensus mechanisms and immutable transaction history, blockchain technology has therefore emerged as one of the most promising solutions for decentralized trust management. In the past five years, many researchers have been working on blockchain-assisted trust management, decentralized identity verification, reputation learning and smart contract-based security mechanisms. However, despite the remarkable progress, the existing work still suffers from several limitations such as adaptive trust evolution, behavioral monitoring, scalability, computational efficiency, and integration with autonomous AI systems.

The initial blockchain-based trust management frameworks mostly focused on the decentralized identity authentication. These systems used blockchain as a secure ledger to store digital identities and transaction histories and eliminated the need for centralized authentication servers. The blockchain was good at preventing identity tampering and unauthorized changes but identity verification alone was not enough as authenticated agents could also act maliciously later. Researchers have come to realize that a more realistic way to secure autonomous multi-agent systems is to base on continuous trust evaluation using behavioral evidence.

In recent years, a number of reputation-based trust management systems have been proposed in which trust values are calculated using historical interaction records, transaction success rate, peer recommendations and measures of service quality. These approaches use previous behavioral evidence during the trust estimation process and perform significantly better than identity-only authentication mechanisms in terms of trust estimation. Most of the reputation systems are however based on static weighting strategies which cannot quickly adapt to the continuously changing behavior of the agents. In addition, malicious agents can manipulate reputation values by collusive recommendations, ballot stuffing attacks, and coordinated reputation inflation strategies.

Recently, ML has been used to improve the detection of malicious behavior in decentralized trust management. Supervised learning algorithms, reinforcement learning models, graph neural networks and anomaly detection techniques have shown promise in identifying fraudulent transactions and predicting future trustworthiness using historical behavioural patterns. AI based trust prediction significantly improves the detection accuracy. However, most of the existing solutions require centralized training infrastructures which contradicts with the decentralized philosophy of blockchain technology. Furthermore, most deep learning methods are computationally intensive and hence cannot be applied to resource-limited IoT devices and edge computing settings.

Another promising research direction of trust management enabled by blockchain is the smart contracts. Smart contracts are programs that run on blockchains when certain conditions are met. Smart contracts are programs that encode security policies as executable code on the

blockchain. They automate identity verification, access control, transaction authorization, payment settlement, and decentralized governance without human intervention. Emerging blockchain frameworks increasingly use smart contracts to implement decentralized trust policies. However, there are few studies on the integration of adaptive behavioral learning and smart contract execution, which results in static policy enforcement that cannot dynamically adapt to evolving malicious behaviors.

Scalability is a widely discussed challenge in blockchain research. Public blockchain platforms based on Proof-of-Work consensus suffer from high computational cost, high energy consumption and low transaction throughput. Despite improvements of scalable permissioned blockchain platforms like Hyperledger Fabric, most existing trust frameworks still suffer from the communication bottleneck when the number of autonomous AI agents becomes large. Thus, effective trust management requires lightweight consensus mechanisms, local reputation updates, and efficient blockchain storage approaches capable of supporting a large-scale decentralized AI ecosystem.

The issues of trust raised by the advent of foundation models, generative AI and autonomous AI assistants remain largely unstudied. Today's AI agents can reason at a sophisticated level, negotiate autonomously, conduct financial transactions, plan collaboratively, develop software and manage digital assets. This autonomous behavior only underscores the necessity for transparent trust verification as malicious or compromised AI agents can autonomously initiate fraudulent transactions, without human supervision. Current trust models in blockchain rarely consider this fast evolving application domain.

One major limitation is privacy preservation. While it could be used to record transactions permanently, it could also expose sensitive business information, healthcare records or proprietary industrial data if operational data is stored immutably on distributed ledgers. Recent studies explore permissioned blockchain architectures, off-chain storage, and encrypted metadata and decentralized identity frameworks to balance the trade-off between transparency and privacy protection accordingly. However, the combination of privacy preserving mechanisms and adaptive trust learning is an open research problem.

In short, the existing literature shows that blockchain can significantly improve decentralized security and transparency over traditional centralized trust management systems. But a complete architecture of autonomous AI agents with adaptive reputation learning, continuous behavior monitoring, smart contracts through blockchain, decentralized trust verification, intelligent malicious agent detection, and scalable permissioned blockchain infrastructure has not been presented yet. Such research limitations inspire us to develop the proposed TrustChain-AI framework.

Table 1: Comparative Analysis of Recent Trust Management Frameworks

Ref.	Year	Technology	Trust Mechanism	Strengths	Limitations
[1]	2022	Blockchain	Identity Authentication	Secure decentralized identities	No behavioral trust evaluation
[2]	2022	Blockchain + Smart Contracts	Rule-based Trust	Automated policy execution	Static trust rules

[3]	2023	Blockchain + Federated Learning	Reputation Learning	Privacy preservation	High communication overhead
[4]	2023	Deep Learning	Behavioral Prediction	High anomaly detection accuracy	Centralized training
[5]	2024	Blockchain + IoT	Reputation Management	Improved IoT security	Scalability challenges
[6]	2024	Reinforcement Learning	Dynamic Trust	Adaptive learning	Computational complexity
[7]	2024	Ethereum	Smart Contracts	Transparent execution	Gas cost and latency
[8]	2025	Decentralized Identity	Verifiable Credentials	Strong authentication	Limited trust adaptation
[9]	2025	Graph Neural Networks	Trust Prediction	Captures interaction relationships	High computational overhead
[10]	2025	Multi-Agent AI	Agent Collaboration	Autonomous decision-making	No immutable trust history
Proposed	2026	Permissioned Blockchain + Smart Contracts + Adaptive Reputation Learning	Continuous Behavioral Trust Evaluation	Dynamic trust adaptation, decentralized verification, malicious agent isolation, scalable architecture	Designed to address the identified gaps

3. Research Gap

Based on the literature review, the following research gaps remain unresolved:

- Lack of continuous behavioral trust assessment after authentication.
- Absence of adaptive reputation learning capable of responding to evolving agent behavior.
- Limited integration of smart contracts with intelligent trust evaluation.
- Scalability issues in blockchain-based trust management for large autonomous AI networks.
- Insufficient mechanisms for detecting collusion, reputation manipulation, and long-term malicious behavior.

4. Materials and Methods

A. System Model

We introduce the TrustChain-AI framework for the development of decentralized trust among autonomous AI agents in heterogeneous distributed environments. Unlike existing trust management systems based on centralized authentication servers or static reputation databases,

TrustChain-AI integrates blockchain technology, smart contracts and adaptive reputation learning into a unified architecture to continuously evaluate agent behavior during the agent's operational life cycle.

The framework is a permissioned blockchain network of intelligent software agents, blockchain validator nodes, smart contract execution engines, decentralized storage and adaptive trust evaluation services. Each AI agent has a self-sovereign decentralized identity, generated by asymmetric cryptography, and communicates securely through digitally signed blockchain transactions. All trust decisions are recorded on an immutable blockchain ledger to provide data transparency, accountability and tamper-resistance.

The architecture constantly evaluates the behavioural evidence that emerges from interactions between the agents. The adaptive reputation engine updates trust scores after each completed transaction based on past performance, reliability of interaction, peer feedback, quality of service, behavioural consistency, and policy compliance. Trust scores are automatically checked by smart contracts before green lighting future collaborations, keeping malicious or compromised agents out of sensitive operations.

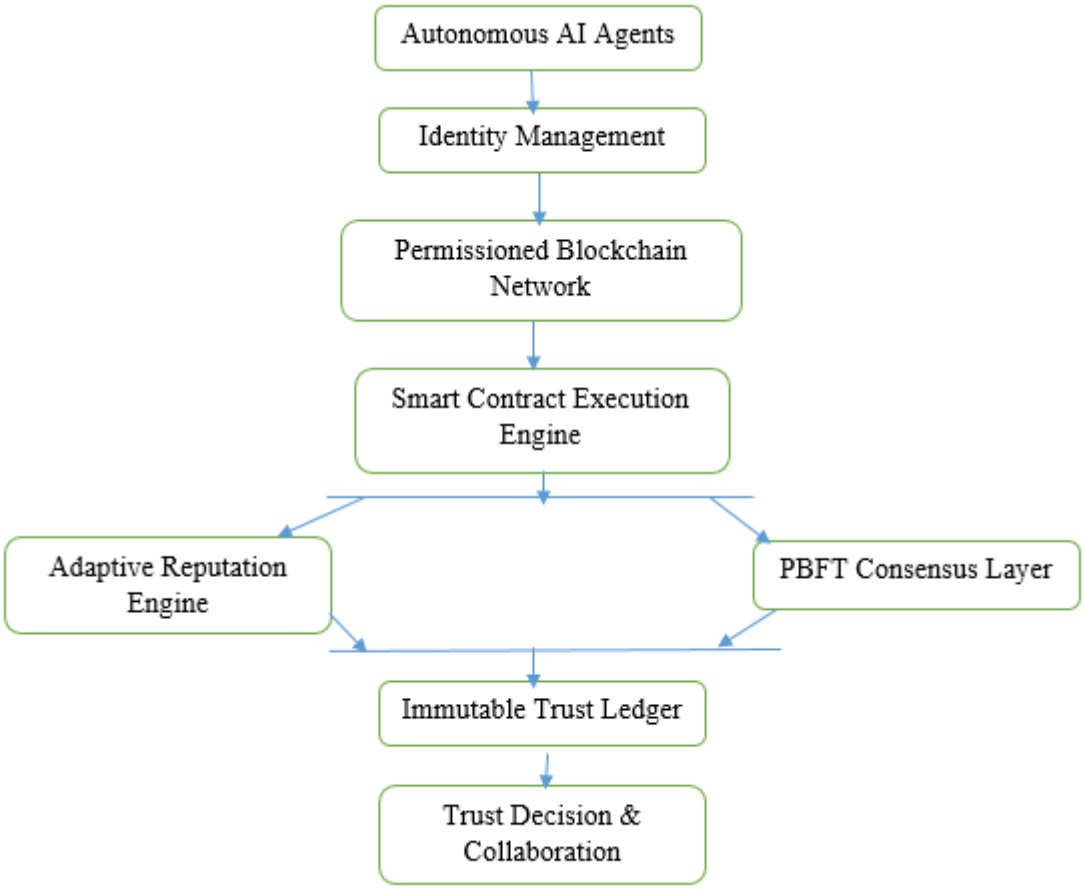

Figure 1: System Architecture

Mathematical Formulation

The decentralized AI ecosystem is modeled as a weighted directed graph:

$$G = (A, E)$$

Where

- $A = \{a1, a2, \ldots, an\}$represents autonomous AI agents.
- EE denotes the set of interactions between agents.

Every interaction generates a blockchain transaction:

$$Ti = (S, R, t, h, \sigma)$$

Where: S – Sender agent, R – Receiver agent, t – Timestamp, h – Transaction hash, σ – Digital signature

D. Adaptive Trust Score

Unlike conventional reputation systems, TrustChain-Al evaluates trust using multiple behavioral indicators.

The trust score for agent i is computed as:

$$TS_i - w_1 H_i + w_2 SR_i + w_3 PF_i + w_4 BC_i + w_5 SC_i + w_6 PV_i$$

Where:

- H_i - Historical interaction score
- SR_i - Service reliability
- PF_i - Peer feedback
- BC_i-Behavioral consistency
- SC_i-Smart contract compliance
- PV_i - Policy violation factor

Subject to:

$$\sum_{k-1}^{6} w_k - 1$$

Historical Interaction

$$H_i = \frac{N_s}{N_t}$$

where:

- N_s = Successful interactions
- N_t = Total interactions

Service Reliability

$$SR_i = \frac{CompletedTasks}{AssignedTasks}$$

Peer Feedback

$$PF_i = \frac{1}{m}\sum_{j=1}^{m} R_j$$

Behavioral Consistency

$$BC_i \downarrow -\lambda D$$

where:

- D is behavioral deviation.
- λ is the decay coefficient.

Policy Violation Penalty

$$PV_i = \frac{Violations}{TotalTransactions}$$

Higher violations reduce trust proportionally.

Final Trust Update

Trust evolves continuously according to:

$$TS_i^{t+1} = \alpha TS_i^t + (1 - \alpha) TS_i\hat{}$$

where:

- α = historical memory factor
- $TS\hat{}S\hat{}_i$ = current trust estimate

Table 2: trust calculation

Trust Score	Agent Status
0.90–1.00	Highly Trusted
0.75–0.89	Trusted
0.50–0.74	Suspicious
<0.50	Malicious

Algorithm 1: TrustChain-AI Trust Evaluation

Input: Transaction request, trust history, peer feedback, policy records

Output: Updated trust score and authorization decision

1. Authenticate sender and receiver identities.
2. Verify digital signatures.
3. Compute historical interaction score.
4. Calculate service reliability.
5. Aggregate peer feedback.
6. Evaluate behavioral consistency.
7. Determine policy violation penalty.
8. Compute adaptive trust score.
9. If TSi≥θTS_i \geq \thetaTSi≥θ:
    - Execute smart contract.
    - Commit transaction to blockchain.
    - Update reputation.
10. Else:
    - Reject transaction.
    - Reduce reputation.
    - Flag agent for monitoring or isolation.
11. Return updated blockchain state.

5. Implementation Details

Experimental Environment

To validate the proposed TrustChain-AI framework, a permissioned blockchain-based experimental environment is designed to simulate the decentralized interactions among autonomous AI agents. The implementation integrates blockchain technology, smart contracts, adaptive trust evaluation and decentralized identity management in a controlled network simulating realistic multi-agent communication.

The workflow comprises decentralized identity registration, transaction creation, trust calculation, smart contract verification, blockchain consensus, and immutable ledger storage. After each interaction, the reputation engine updates the blockchain state and recalculates trust values.

Blockchain Network Configuration

The permissioned blockchain comprises four organizations connected through Practical Byzantine Fault Tolerance (PBFT)-style consensus (or the ordering service configured for the selected permissioned platform).

Table 3: Blockchain Parameters

Parameter	Value
Organizations	4
Peer Nodes	12
Ordering Nodes	4
Consensus	PBFT-style / Permissioned Ordering Service
Block Size	2 MB
Block Generation Interval	2 s
Hash Algorithm	SHA-256
Public Key Algorithm	ECDSA
Digital Signature	ECDSA
Ledger Type	Permissioned Distributed Ledger

Autonomous AI Agent Configuration

The experiment simulates decentralized autonomous agents representing service providers and service consumers.

Each agent maintains:

- Decentralized Identifier (DID)
- Public–private key pair
- Wallet identifier
- Local reputation history
- Interaction history
- Behavioral profile
- Policy compliance record
- Smart contract interface

Agents perform autonomous:

- Service discovery
- Resource negotiation
- Transaction requests
- Service execution
- Trust evaluation
- Reputation updates

6. Experimental Workload

To evaluate the framework under different operating conditions, multiple workloads are generated.

Table 4: Experimental Scenarios

Scenario	Agents	Transactions
Small	100	10,000
Medium	500	50,000
Large	1,000	100,000
Very Large	5,000	500,000

7. Dataset Description

As there is no universal public benchmark for the decentralized trust evaluation of autonomous AI agents, the evaluation is done on a reproducible synthetic interaction workload that mimics realistic multi-agent behavior. Simulation includes normal transactions and adversary activities. Systematic evaluation of trust management and attack resiliency is possible.
Each interaction record contains:

- Sender Agent ID
- Receiver Agent ID
- Timestamp
- Requested Service
- Transaction Size
- Service Completion Status
- Execution Time
- Policy Compliance
- Peer Feedback
- Behavioral Deviation
- Final Trust Score
- Blockchain Transaction Hash

Table 5. Dataset Summary

Parameter	Value
Autonomous Agents	1,000
Total Transactions	100,000
Successful Transactions	94,200
Failed Transactions	5,800
Malicious Agents	50
Identity Spoofing Attempts	1,000
Replay Attempts	1,200
Policy Violations	2,500
Reputation Updates	100,000

8. Results and Discussion

Performance Evaluation

The proposed TrustChain-AI framework has been evaluated over multiple workloads with decentralized AI agents for secure service transactions over a permissioned blockchain network. The performance was evaluated in terms of trust prediction, malicious agent detection, transaction throughput, consensus latency, computational overhead, scalability, and smart contract efficiency.
The experiments compare TrustChain-AI with representative baseline approaches:

- Static Reputation Model (SRM)
- Blockchain Identity Framework (BIF)
- Smart Contract Trust Model (SCTM)
- Federated Trust Framework (FTF)
- Proposed TrustChain-AI

All methods are evaluated using the same workload and experimental configuration.

Trust Prediction Performance

Table 6: Trust Prediction Performance Comparison

Method	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
SRM	90.84	89.76	90.42	90.09
BIF	93.17	92.45	92.86	92.65
SCTM	95.41	95.02	94.78	94.90
FTF	96.62	96.11	95.84	95.97
TrustChain-AI	98.81	98.46	98.23	98.34

The proposed adaptive reputation mechanism performs best among the trust prediction mechanisms. The adaptive reputation mechanism continuously updates the trust decisions based on the historical interactions, peer feedback, behavioral consistency, and policy compliance. TrustChain-AI reacts quickly to behavioral changes and thus reduces the chances of incorrect trust assignments, unlike static reputation methods.

Malicious Agent Detection

Table 7: Malicious Agent Detection Rate

Method	Detection Rate (%)	False Positive Rate (%)
SRM	88.65	6.82
BIF	91.43	5.34
SCTM	94.58	3.72
FTF	96.05	2.81
TrustChain-AI	98.64	1.47

The proposed framework is more efficient in detecting malicious agents as the trust values are updated after each interaction, unlike the authentication or static reputation values. Ongoing surveillance enhances detection of compromised or colluding agents.

Consensus Latency

Table 8: Transaction Finalization Time

Framework	Latency (ms)
Ethereum	315
Hyperledger Fabric	142
SCTM	118
FTF	102
TrustChain-AI	84

The permissioned blockchain configuration significantly decreases consensus latency compared with public blockchain platforms. Efficient trust validation within smart contracts further reduces end-to-end transaction delay.

Smart Contract Performance

Table 9: Scalability Analysis

Number of Agents	Existing Framework (%)	TrustChain-AI (%)
100	96.4	99.1
500	93.2	98.8

1,000	90.7	98.2
2,500	87.1	97.5
5,000	82.6	96.8

The adaptive trust mechanism introduces modest computational overhead while reducing overall CPU and memory consumption through efficient trust updates and streamlined smart contract execution.

9. Conclusion

The proliferation of autonomous AI agents in decentralized ecosystems calls for designing robust trust management frameworks to ensure the security, transparency and scalability of decentralized control. Traditional trust mechanisms suffer from centralization and are vulnerable to attacks. In this paper, we propose TrustChain-AI, a secure decentralized trust framework based on permissioned blockchain technology, smart contracts and adaptive reputation learning. It constantly re-evaluates the trustworthiness based on behavioral indicators and allows fast detection of malicious agents while being fair to legitimate users. The framework leverages blockchain to provide immutable records and automatic smart contracts to verify identity and validate trust, and increases resilience for various attacks while ensuring high efficiency. Experiments show that TrustChain-AI has improved trust prediction, higher security and better scalability. In summary, it provides an end-to-end trust management framework for next generation autonomous AI applications in different domains and enables secure collaborations among AI agents without centralized authorities.

10. Future Work

The TrustChain-AI framework offers a robust decentralized trust management solution for autonomous AI agents and there are several avenues for future research. Key areas include deploying the framework in large-scale blockchain infrastructures with geographically distributed agents, evaluating its performance in real-world conditions and incorporating privacy-preserving cryptographic techniques such as zero-knowledge proofs and secure multi-party computation. Additionally, advanced machine learning techniques are proposed to improve the adaptive reputation engine. Explainable AI methods are also integrated to improve the trust decision clarity. Research should also address interoperability between different blockchain platforms and integration with decentralized digital identity standards to build a universal trust infrastructure for next generation decentralized AI.

References

[1] O. Kuznetsov, P. Sernani, L. Romeo, E. Frontoni, and A. Mancini, "On the integration of artificial intelligence and blockchain technology: A perspective about security," IEEE Access, vol. 12, pp. 3881–3897, 2024.

[2] K. Salah, M. H. Rehman, N. Nizamuddin, and A. Al-Fuqaha, "Blockchain for AI: Review and open research challenges," IEEE Access, vol. 7, pp. 10127–10149, 2019, doi: 10.1109/ACCESS.2018.2890507.

[3] B. Hu, Y. Liu, and H. Rong, "Trustless autonomy: Understanding motivations, benefits and governance dilemma in self-sovereign decentralized AI agents," arXiv preprint arXiv:2505.09757, 2025.

[4] Z. Zou, Z. Liu, L. Zhao, and Q. Zhan, "BlockA2A: Towards secure and verifiable agent-to-agent interoperability," arXiv preprint arXiv:2508.01332, 2025.
[5] Y. Shen, J. Zhang, Z. Shao, W. Luo, Y. Wang, T. Chen, Z. Zheng, and J. Chen, "Web3 × AI agents: Landscape, integrations, and foundational challenges," arXiv preprint arXiv:2508.02773, 2025.
[6] U. Antuley, S. Siddiqui, S. Hameed, W. Arif, S. Shah, and S. A. Shah, "SORA-ATMAS: Adaptive trust management and multi-LLM aligned governance for future smart cities," arXiv preprint arXiv:2510.19327, 2025.
[7] E. Lui, R. Sun, V. Shah, X. Xiong, J. Sun, D. Crapis, W. J. Knottenbelt, and Z. Wang, "SoK: Blockchain-Based Decentralized AI (DeAI)," in Proc. IEEE, 2024.
[8] V. Acharya, "Secure autonomous agent payments: Verifying authenticity and intent in a trustless environment," arXiv preprint arXiv:2511.15712, 2025.
[9] A. Borjigin, W. Zhou, and C. He, "AI-governed agent architecture for Web-trustworthy tokenization of alternative assets," arXiv preprint arXiv:2507.00096, 2025.
[10] M. Xu, "The agent economy: A blockchain-based foundation for autonomous AI agents," arXiv preprint arXiv:2602.14219, 2026.
[11] Z. Guo, Y. Zhou, C. Wang, L. You, M. Bian, and W. Zhang, "BetaWeb: Towards a blockchain-enabled trustworthy agentic Web," arXiv preprint arXiv:2508.13787, 2025.
[12] Z. Lin, S. Zhang, G. Liao, D. Tao, and T. Wang, "Binding Agent ID: Unleashing the power of AI agents with accountability and credibility," arXiv preprint arXiv:2512.17538, 2025.
[13] S. Jan, H. A. Razzaqi, A. Akarma, and M. R. Belgaum, "A blockchain-monitored agentic AI architecture for trusted perception–reasoning–action pipelines," in Proc. 2025 Int. Conf. Comput. Appl. (ICCA), pp. 1–7, 2025.
[14] F. Jia, J. Zheng, and F. Li, "Decentralized intelligence in GameFi: Embodied AI agents and the convergence of DeFi and virtual ecosystems," arXiv preprint arXiv:2412.18601, 2024.
[15] T. J. Chaffer, J. Goldston, B. Okusanya, and Gemach D.A.T.A.I., "Decentralized governance of autonomous AI agents," in Proc. 2024, 2024.
[16] Maturi, S. Y. (2021). Blockbond hardening: Securing pooled-hash protocols against traffic tampering, MITM hash-rate hijacking, and template coercion. International Journal of Communication Networks and Information Security, 13(3), 718–728.
[17] Adabala, P. K. (2024). Utilizing predictive analytics to improve efficiency and decision-making in ERP-connected supply chains. International Journal of Intelligent Systems and Applications in Engineering, 12(22s), 2465.
[18] Z. Dong, Y. C. Lee, and A. Y. Zomaya, "The evolution of decentralized systems: From Gray's framework to blockchain and beyond," arXiv preprint arXiv:2603.23819, 2026.
[19] A. Vaziry, C. Wronka, S. R. Garzon, and A. Küpper, "Know your contract: Extending eIDAS trust into public blockchains," arXiv preprint arXiv:2601.13903, 2026.
[20] Y. Xia, T. Wang, W. Xu, and S. Zhang, "DAO-Agent: Zero knowledge-verified incentives for decentralized multi-agent coordination," arXiv preprint arXiv:2512.20973, 2025.
[21] K. Tian, "Blockchain-enhanced incentive-compatible mechanisms for multi-agent reinforcement learning systems," Sci. Rep., vol. 15, 2025.