A Validatable Multi-Agent Collaboration Mechanism Driven by Smart Contracts

Sangtian Guan, Fei Lin, Juanjuan Li, and Xiaolong Liang

***Abstract*—Artificial intelligence (AI) agents driven by large language models have demonstrated substantial advancements in addressing complex problem-solving tasks. While current AI agents' collaborations are predominantly organized by centralized platforms, this organizational model inherently introduces systemic challenges such as data monopolies and interoperability barriers, thereby hindering cross-domain collaborative potential. Although current multi-agent collaboration methods have enabled secure data and resource sharing, they still fall short in utilizing distributed knowledge and adopting validatable collaborative workflows. To overcome these limitations, this paper proposes a distributed multi-agent collaboration mechanism leveraging the smart-contracts-based validation method with specifically designed planning, execution and validation smart contracts for robust workflow validation. An illustrative example of decentralized science collaborative writing is provided to demonstrate the applicability of the proposed mechanism. To further bridge the data silos and autonomously update the smart contracts, we employ decentralized autonomous organizations and operations (DAOs) and adopt an auction-based tool supplier matching model to architect the organizing and incentive scheme to enhance self-adaptivity and scalability. To further evaluate the performance of the proposed mechanism, we delineate a hierarchical indicator system and discuss the corresponding key technologies. This paper outlines a smart-contract-based approach to enhance the reliability of multi-agent collaboration, which could serve as a methodological basis for constructing future distributed multi-agent collaboration ecosystems.**

I. INTRODUCTION

Artificial intelligence (AI) agents are intelligent entities powered by advanced large language models (LLMs), whose performance could surpass the human experts in several areas [1]. AI agents typically integrate several essential abilities such as planning, tool-using, memory, and contextual awareness based on their foundational model that enables the goal-driven behaviors and dynamic adaptation capabilities [2], aiming to address complex multimodal and multidomain problems [3, 4].

However, the capability of a single agent suffers from interoperability constraints of tool-using stemming from divergent training methodologies of foundation models, exacerbating the capacity to handle cross-domain, multimodal, or complex reasoning tasks [5]. Multi-agent collaboration is a trending solution to tackle this problem by enhancing the representative ability of domain knowledge and professional skills [6]. Current multi-agent systems are typically deployed via specific platforms that monetize internal transaction volume [7]. However, this paradigm perpetuates data monopolies across platforms and creates data silos among users [8], thereby restricting the richness and diversity of domain knowledge available for AI agent collaboration [9].

To address this challenge, existing studies have adopted blockchain technologies to bridge personal data silos and break data monopolies [10, 11]. With the development of Web 3.0, recent advancements in decentralized AI frameworks have demonstrated the potential of blockchain-based marketplaces to facilitate collaborative machine learning by enabling secure data sharing and distributed model training [12]. However, on the one hand, the current works fall short in enabling knowledge extraction from data sharing and collaboration, and thus hinder the orchestration of collaborative problem-solving. On the other hand, multi-agent collaboration frameworks, such as MetaGPT [13] and Fincon [14], could enhance the robustness, adaptability, and generalizability on diverse tasks by planning a workflow and communication to process specified subtasks and share resources [6]. However, the ignorance of the workflow validation mechanism makes the cumulative error propagation and state inconsistency [15], further leading to an inevitable failure. To tackle this challenge, SagaLLM [16] proposed a transaction-based coordination mechanism validation, and AgentCoord [17] designed a visual exploration framework to facilitate and validate the development of coordination strategies. Nevertheless, these solutions lack the flexibility to adapt to distributed environments where data silos have been partially mitigated, which demand both data interoperability and dynamic collaborative workflows.

By virtue of the blockchain foundation, smart contracts inherently possess automatic execution with the traceability

Manuscript received: 27 October 2025; revised: 20 November 2025; accepted: 26 November 2025. *(Corresponding author: Juanjuan Li.)*

Citation: S. Guan, F. Lin, J. Li, and X. Liang, A validatable multi-agent collaboration mechanism driven by smart contracts, *Int. J. Intell. Control Syst.*, 2025, 30(4), 349–356.

Sangtian Guan and Fei Lin are with Faculty of Innovation Engineering, Macau University of Science and Technology, Macao 999078, China (e-mail: sangtian.guan@ieee.org; feilin@ieee.org).

Juanjuan Li and Xiaolong Liang are with State Key Laboratory of Multimodal Artificial Intelligence Systems, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China (e-mail: juanjuan.li@ia.ac.cn; xiaolong.liang@ia.ac.cn).

Digital Object Identifier 10.62678/IJICS202512.10261

and immutability of the state dynamic [18], further characterized by functional programmability and executional atomicity [19]. Furthermore, smart contracts can guarantee the executional determinism and operational reliability of distributed multi-agent collaborations, thus facilitating the coherent execution of complex collaborative workflows. On this basis, the smart contracts can support the organization, collaboration, and execution of the decentralized autonomous organizations and operations (DAOs) for multi-agent collaboration [20], the main organizational infrastructure of the Web 3.0, where community members' goals can be aligned to get involved in a scalable and self-adaptive collaborating paradigm.

Based on these considerations, this paper proposes a smart-contract-based distributed agent collaboration mechanism to derive an atomic validation with its corresponding management and incentive mechanisms for an open and dynamic collaborative ecosystem. Firstly, we propose a smart-contract-based validation mechanism for distributed multi-agent collaboration workflows and an illustrative example on decentralized science (DeSci) collaboration, aiming to enhance the state consistency and overall reliability in collaborative executions. Secondly, we demonstrate a DAO-based incentive mechanism that incentivizes high-quality tool availability and expert-curated knowledge sharing with an auction and autonomous governance mechanisms, promoting the self-adaptivity and scalability of the collaborating system. Thirdly, we discuss the corresponding indicator systems and the key technologies with the proposed mechanism for further benchmarking and optimizing the proposed mechanism.

The rest of this paper is organized as follows. Section II introduces the basic components, Section III brings the proposed collaboration mechanism, Section IV illustrates the incentive mechanism for the mechanism participants, Section V discusses the indicator system and key technologies, and Section VI gives a conclusion for this paper.

II. Basic Components

A. Participants

We categorize participants into three types: requesters initiate tasks, tool suppliers provide necessary services, and domain experts contribute specialized expert-curated knowledge, aiming at facilitating structured task execution.

Requesters are the task initiators, specifically issuing tasks and providing signed authentication for each request. The requester publishes clearly defined objectives and detailed requirements to ensure that all collaborative agents have an accurate, validated foundation for execution.

Tool suppliers offer specialized tools utilizing services, required to append their digital signatures to all returned results. The signature maintains the alignment between tool invocation and contract execution, safeguarding execution integrity and thereby fostering robust and reliable collaboration within the distributed multi-agent environment. This role ensures reliable data provisioning, supporting robust interactions among participants.

Domain experts contribute essential specialized knowledge and manage their private AI agents that access expert-curated knowledge repositories locally, and are required to append the digital signatures together with the agents' outputs. The domain experts' participation is critical for embedding authoritative, expert-curated knowledge into the collaborative process, significantly enhancing the accuracy and reliability of collaborative outcomes.

B. Specified Smart Contracts

The validatability of the proposed mechanism is supported by three specialized smart contracts: the planning, execution, and validation contracts. These contracts facilitate structured, transparent, and reliable interactions among AI agents. The smart contracts could be deployed in Ethereum virtual machine (EVM)-compatible blockchain and implemented in standard primitives with Solidity, RUST, or Move, so the mechanism remains largely blockchain-agnostic and can be ported to other platforms with equivalent capabilities.

(1) Planning smart contracts establish a predetermined schema for task planning, serving as the foundational workflow for collaborative AI agents. Specifically, these contracts ingest structured planning inputs to outline task execution sequences systematically. Each task segment within the workflow is annotated with detailed metadata, including required knowledge domain labels, the corresponding execution tools, and inter-task dependencies to ensure clarity and completeness in execution pathways, thereby facilitating precise resource allocation and clarifying the operating state for potential rollback in case of any subtask failure. To ensure the standardization and machine-readability of information extracted during task planning, a structured output specification mechanism is integrated into the planning smart contracts as a task-agnostic template. Specifically, standardized output templates are designed to enforce fixed-format outputs from the AI agents. These templates support multiple data exchange formats, such as JSON, XML, or YAML, for the structured representation of essential elements, including task descriptions, temporal constraints, required resources, and interdependencies [21, 22]. This guarantees uniform structure across all planning outputs, facilitating subsequent processing, validation, and analysis.

These procedures could enhance the usability and normative quality of the generated plans. Planning smart contracts incorporate intrinsic integrity assertions to validate the output against predefined syntactic and semantic specifications, to ensure that the generated planning outputs meet the operational and collaborative format requirements of the multi-agent system, thereby enhancing collaborative efficiency by minimizing ambiguity in agent coordination.

(2) Execution smart contracts function primarily as the validation mechanism to ensure that the domain experts with the correct labels assigned and the tool requests issued by the experts' AI agents are executed, instead of a hallucination. For each subtask, an execution contract will be instantiated that binds the subtask to a designated domain expert and to the set of tools that this expert may legitimately invoke. Executing smart contracts contain a dual-stage command flow. Firstly, executing the smart contracts will take the AI agent's query as

input and return it with its corresponding tool or the expert's processed result. During this stage, upon receiving a tool request, the contract initiates a matching process that identifies a qualified tool supplier capable of fulfilling the requested service, in which the exact process will be detailed in Section IV. Secondly, executing the smart contracts will take the tool supplier's digital signature as input, and pass a "finish" state to the blockchain, indicating AI's function calling has been done with the assigned tools.

This contract ensures accountability and adherence to the planned workflow, thus potentially enhancing reliability and consistency across multi-agent interactions.

(3) Validation smart contracts are designed to ensure compliance of agent outputs with predefined formatting standards, directly aligning outcomes with initial task specifications. Firstly, the validation smart contracts will be activated only if there is an observable "finish" state of its prefix contract in the blockchain. When activated, the validation smart contracts will take the AI agent's output content to check whether this matches the predefined format. By employing "if-then" conditional validations, validation smart contracts will scrutinize key parameters such as text length or grammar structure. After that, the validation smart contracts need the digital signature of the expert to output a "finish" state to the blockchain to ensure the atomicity of the retrieval of the expert-curated knowledge. Any deviations trigger immediate feedback loops, mandating corrective revisions by the involved agents.

Through such stringent validation measures, the validation of smart contracts effectively standardizes outputs, thereby maintaining uniformity, coherence, and quality across multi-agent collaborative outcomes.

III. Validatable Multi-Agent Collaboration Mechanism

This section brings the proposed validatable multi-agent collaboration mechanism that orchestrates task planning, tool-enabled execution, and expert validation through corresponding smart contracts, illustrated by a distributed collaborative writing scenario.

A. Proposed Mechanism

The collaboration mechanism initiates when a requester issues a prompt clearly detailing the task requirements, as shown in Fig. 1. Subsequently, the requester's local AI agent performs comprehensive task planning, structuring the task into a hierarchical tree composed of interdependent subtasks, which aligns closely with the original requester's intent, clearly indicating the required domain-specific expert labels, the necessary execution tools, and specific output formatting requirements.

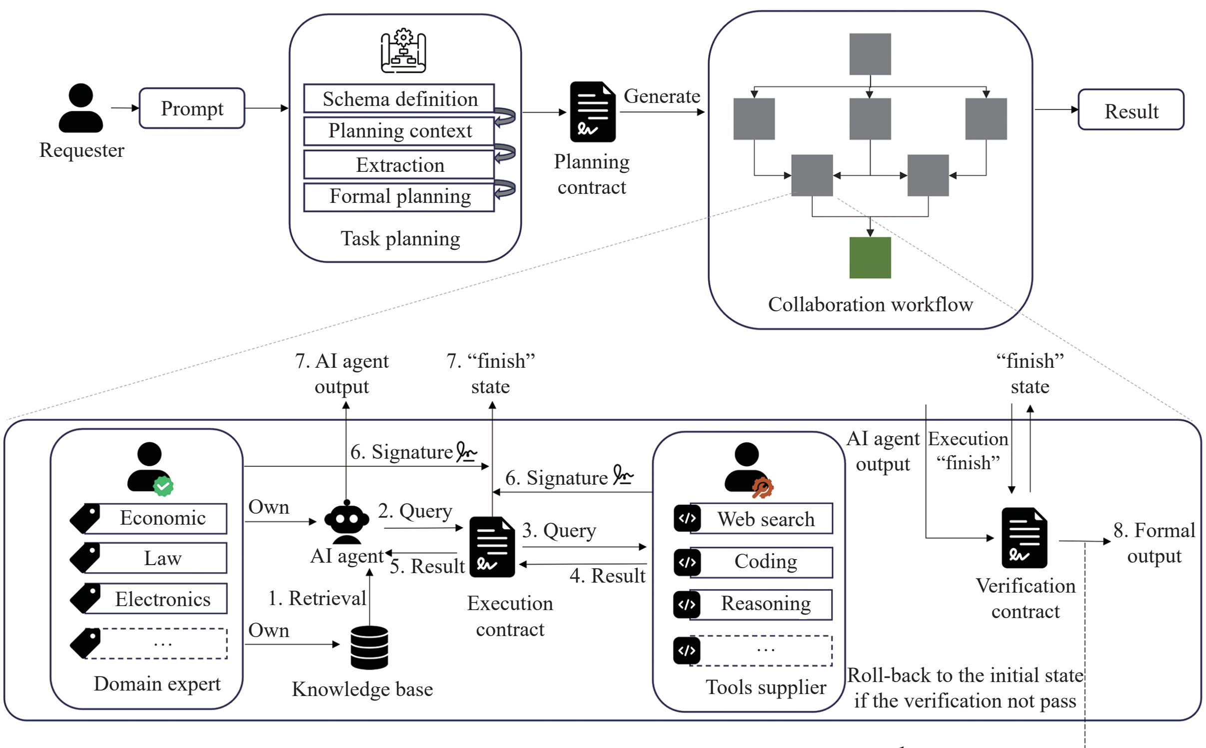

Fig. 1 The multi-agent collaboration mechanism based on smart contracts.

Once this structured plan outputs, it serves as input for the planning smart contract and generates specific execution and validation smart contracts corresponding to the subtasks defined. Each subtask is allocated to a domain expert's private AI agent associated with the relevant expert labels. AI agent proceeds by invoking contract-defend tools, simultaneously documenting invocation details through the execution smart contracts.

In the action phase, according to prescribed subtask formats, the AI agent will consult expert-curated knowledge repositories, collecting the tool returning and retrieving expert-curated knowledge to enrich its prompting context. Output will be submitted to its corresponding validation smart contract, which first validates whether the "finish" status is logged by the execution contract, then the digital signatures of the corresponding experts and suppliers. These signatures certify the inclusion and proof of the resources of the knowledge base and tool-using. After that, validation contracts evaluate the compliance of the outputs against specified formatting standards such as text length, structure, and designated encoding formats. Only after successfully passing these validations does the validation smart contracts issue the "finish" status update to the blockchain. If compliance fails at any stage, the contract triggers corrective actions, mandating output revisions until all criteria are satisfied. Once an AI agent fails or hallucinately invokes the specified smart contracts pertinent to its assigned tasks, collaborating nodes will trigger a rollback mechanism, compelling the agent to restart and correctly execute the designated procedures. This iterative process continues systematically until the final task is fulfilled and the complete outputs are delivered.

Within this mechanism, to reduce on-chain storage and gas costs while preserving verifiable audit trails for critical coordination and accountability events, participants' interaction histories and the execution states of planning, execution, and validation contracts, including task identifiers, signatures, and status transitions, are recorded on-chain. Computationally intensive and data-heavy resources, such as expert knowledge bases, AI agents, and external tools, remain off-chain and are accessed through authenticated calls.

Several key differences between conventional centralized multi-agent collaboration methods and the proposed smart-contract-based mechanism are qualitatively summarized in Table 1.

Table 1 Comparison of feature characteristics with the conventional methods.

Feature dimension	Conventional methods	Proposed mechanism
Participant roles	Users must join a single platform and act within its fixed role model and APIs.	Requesters, domain experts, and tool suppliers can join as independent participants.
Collaborations	AI agents are typically hosted on a centralized server controlled by the platform provider, where participants must connect to this single orchestration point.	Allowing geographically distributed participants to collaborate through their local AI agents, tools, and infrastructure.
Services	Restricted to internal toolsets or specific APIs defined by the framework developers.	Cross-domain tool usage and expert-curated knowledge can be shared through standardized interfaces and metadata, enabling composition and reuse across heterogeneous systems.
Execution reliability	Execution logic is implemented off-chain in opaque services, prone to "hallucinated" tool usage where the agent claims action without execution.	The smart contracts cryptographically guarantees that tool invocation and function calling actually occurred.
State consistency	Local, mutable, and lacking by default cross-organization verifiability and immutability.	Leverages the EVM to ensure state consistency, traceability, and immutability of the collaboration workflow.

B. Illustrative Example

To illustrate the proposed mechanism, as shown in Fig. 2, we consider a distributed collaborative writing scenario of DeSci collaborative writing [23], which is one of the most important components of the Web 3.0 and necessitates clear specifications regarding textual division, word count, and format, alongside cross-domain expert validation for accuracy. The process initiates when a requester submits a prompt outlining the desired topic entry. This prompt is subsequently received by a task-planning expert agent, which utilizes historical data to formulate an effective collaborative writing plan. This detailed collaborative workflow is formalized and fed into the planning smart contracts.

Upon receiving the formalized plan, the planning smart contracts generate interrelated execution and validation smart contracts. Execution contracts, determined through a competitive bidding process, are allocated to appropriate tool suppliers responsible to ensure the atomicity of tool executions. In parallel, validation contracts are assigned to respective domain experts tasked with validating the compliance of output content with specified textual length and formatting requirements (such as LaTeX or Markdown formats).

Once all contracts are executed successfully, the final results will be consolidated and returned to the requester to ensure comprehensive and accurate fulfillment of the initial prompt and collaboration format specifications.

IV. Organization and Incentive Scheme

This section introduces the DAO-based collaborating incentive mechanism, which is composed of three specified models aimed respectively at dynamically matching suitable suppliers to functional calls, maintaining consistency between contract execution, function calling, and expert-curated knowledge retrieval, and demonstrating the openness and self-optimization capabilities of the proposed system.

A. Multi-Agent Collaboration DAO

To tackle the issues brought by the centralized platform while overcoming the static vulnerable challenges of smart contract, DAOs should be integrated as a dominant paradigm of member managing, collaborating system maintenance, and contracts update [24]. DAO's autonomous and transparent governance mechanism introduces scalable functionalities and mediates conflicts arising from insufficient quality provided by tool suppliers and domain experts. Community members, including requesters, suppliers, and experts, may initiate governance proposals addressing new feature requests or lodging complaints against underperforming participants. Proposals are subject to community scrutiny through a transparent voting process to ensure democratic decision-making. Approved proposals result in system-wide implementation of new features or the enforcement of corrective actions, including restrictions or expulsions of non-compliant suppliers or experts. By utilizing blockchain as the decentralized ledger

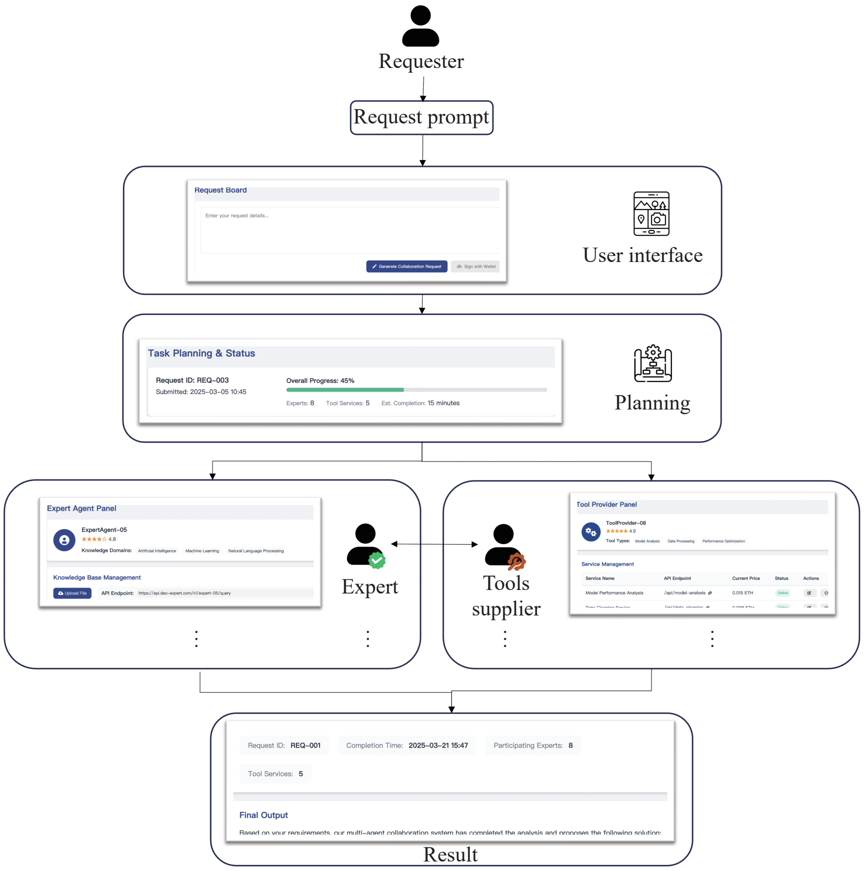

Fig. 2 Distributed collaborative writing system and its user interface.

and data infrastructure, each validated interaction updates participants' on-chain anonymous histories, enabling reputation scores and access control policies that are transparent and tamper-resistant. Decentralized identities (DIDs) are linked to verifiable histories of the community members, where malicious suppliers and Byzantine agents are discouraged through slashing, access restrictions, and negative reputation updates triggered by governance proposals. Sybil attacks are constrained by identity binding and reputation-weighted voting. On this basis, a stablecoin-driven economic system can be introduced to construct a relatively stable on-chain economic environment that decouples task economics from crypto-asset volatility. Requesters pay for successful subtasks in stablecoins, while suppliers and experts receive performance-contingent payments, refunds, or penalties directly via smart contracts. This mechanism fosters continuous evolution, enforces service quality, improves the accountability for cost-benefit analysis, and enhances the ecosystem's robustness against strategic manipulation.

Tool suppliers offering the same service type compete for provision rights via a sealed second-price reverse auction, wherein the lowest bidder wins but pays the second-lowest bid price. To enhance the distribution fairness and cost-efficiency for the service offering, we integrate an auction-based tool supplier bidding model.

B. Auction-Based Tool Supply-Demand Matching Model

Since AI agents could serve as decision assistants, we assume that suppliers in collaboration with DAOs exhibit individual rationality by leveraging AI-aided decisions to pursue utility maximization. Meanwhile, the transparency and immutability of blockchain ensure that all participants share a consistently distributed ledger, theoretically eliminating information asymmetry. Thus, we can consider the supply bidding game with complete information. Besides, every tool supplier holds independent private value for the heterogeneity of its private hardware and software asset status. Consider a set of suppliers $S = \{s_1, s_2, \dots, s_n\}$, where each supplier s_i has an independent true valuation v_i for providing a specific tool service and submits a sealed bid b_i. Each supplier s_i confidentially submits a bid $b_i \geqslant 0$, forming a bid vector $B = (b_1, b_2, \dots, b_n)$. The winning supplier s_w is determined as the bidder with the highest submitted bid $w = \arg\max_{i \in S} b_i$.

Supplier s_w obtains the right to provide the service at a price equal to the second-highest bid $p = \max_{i \in S \setminus \{w\}} b_i$. The utility for each supplier s_i is thus defined by

$$u_i = \begin{cases} v_i - p, \ i = w; \\ 0, \ i \neq w. \end{cases}$$

The dominant strategy equilibrium occurs when each rational supplier truthfully submits a bid equal to its true valuation $b_i^* = v_i$. This mechanism satisfies dominant-strategy incentive compatibility to ensure that truthful bidding is the dominant strategy. Additionally, a subsequent community-based voting mechanism complements the auction, providing an incentive for the winning supplier to maintain high-quality service to avoid the negative utility associated with expulsion from the system [25]. This mechanism is incentive compatible, inducing strategic truthfulness as each supplier's dominant strategy is to bid its true valuation, since misreporting yields no advantage under the second-price rule [26]. By eliciting honest bids, the auction efficiently allocates the service contract to the supplier best positioned to deliver high quality, for the bidder with the genuinely highest value or lowest cost. To ensure ongoing performance, this auction is complemented by a community voting protocol that permits the expulsion of underperforming suppliers. The looming threat of collective expulsion functions as a post-selection incentive for the winning provider to maintain effort and service standards, guarding against the moral hazard of "bid low, then shirk" behavior [27]. The truth-enforcing auction and community oversight create a mechanism that selects a high-value provider and motivates it to uphold service quality, thereby deterring low-effort performance.

V. Discussion

Considering the empirical utilization for the future multi-agent systems and the potential for continuous improvement, we discuss the corresponding hierarchical evaluation critical system with three key technologies.

A. Evaluation Indicator System

a. Contract-Level Semantic Alignment

How faithfully on-chain predicates encode real-world task requirements and how well these encoded rules cover realistic outputs are essential questions to determine how meaningful contract-level verification is in practice. First, semantic faithfulness can be examined by aligning natural-language task specifications and expert-designed templates with the corresponding on-chain if-then predicates, and by conducting expert reviews or annotation-based studies to assess whether the encoded conditions correctly reflect the intended constraints on subtasks, to clarify where contractual abstractions may omit important aspects of the real-world tasks. Second, rule coverage and robustness can be empirically evaluated using benchmark suites of valid and invalid outputs for different task types, and by measuring the precision, recall, and false positive or negative rates of contract-based format and workflow checks, to provide a quantitative view of how well the smart-contract validators bridge the semantic gap between high-level task requirements and low-level execution traces.

b. Organization-Level Efficiency

At the organizational layer, the proposed mechanism should improve both collaborative efficiency and economic cost efficiency relative to centralized platforms. First, collaborative efficiency can be measured on matched tasks by comparing end-to-end completion time, the number of failed or ambiguous intermediate outputs, and the number of human review iterations required to reach an acceptable draft across baseline centralized workflows and the proposed mechanism, to reflect how effectively the standardized planning and validation pipeline reduce coordination overhead. Second, economic cost efficiency can be quantified by computing the requester's expected payment for successfully validated tool services under the auction-based model and contrasting it with fixed-price procurement on centralized platforms with comparable service quality. Additional indicators, such as the variability of supplier revenue and the share of payments associated with rolled-back or invalid subtasks, can further illuminate the economic impact of the mechanism.

c. Coordination-Level Reliability

Reliability depends on both local error rates at individual tool calls and the way errors accumulate or are corrected along collaborative workflows. First, local error behavior can be evaluated by logging hallucinated or ungrounded tool invocations, invalid intermediate outputs before validation, and the fraction of tool calls that are intercepted by the execution and validation contracts. Comparing these statistics with those from baseline centralized orchestrations shows how effectively the mechanism constrains unverified AI agent actions. Second, global error propagation can be assessed via coordination-level indicators such as the proportion of subtasks that require rollback, the number of corrective cycles per task, and the overall task failure or abandonment rate. Controlled experiments on matched workflows under alternative coordination methods and the proposed design can then quantify the extent to which cumulative errors are reduced.

d. System-Level Performance

At the system level, the viability of the mechanism hinges on workflow scalability and blockchain resource efficiency, which jointly determine whether large-scale, geographically distributed deployments are feasible. First, scalability can be evaluated by systematically increasing the number of agents, subtasks, and concurrent workflows, while measuring end-to-end latency, task throughput, and chain utilization. Stress tests can quantify how performance degrades as the collaborative environment grows. Second, blockchain resource efficiency can be assessed by recording gas consumption and on-chain storage overhead per subtask and per workflow, and by decomposing these costs across planning, execution, and validation contracts. This enables a principled analysis of the trade-off between stronger validation guarantees and additional on-chain cost.

B. Key Technologies

a. Privacy-Preserving Reasoning Algorithm

Personal privacy-preserving is a pivotal factor that demonstrates the trustworthiness of users in various scenarios of Web 3.0 [28]. Employing a privacy-preserving reasoning algorithm is motivated by the imperative of safeguarding sensitive input and output data. Utilizing technologies such as zero-knowledge proofs, this approach would enable encrypted yet verifiable information exchanges, significantly enhancing data privacy. Such privacy enhancement ensures broader applicability and user acceptance of the collaborative mechanism by addressing critical confidentiality concerns.

b. Trajectory-Based Reputation Mechanism

Considering the dynamic role of AI agents in fostering interdisciplinary innovation, traditional centralized identity validation authorities face inherent limitations due to rigid compliance management procedures, making them ill-suited for rapidly evolving and highly specialized domains [29]. In this context, a decentralized, community-driven identity validation system becomes increasingly crucial. Leveraging blockchain-stored historical interaction data and self-selected identity labels, combined with AI agent-driven dynamic analytical capabilities, this mechanism continuously evaluates and updates member reputations. This adaptive reputation profiling not only enhances the accurate identification and characterization of highly specialized domain experts but also enables more personalized and efficient collaboration among agents. Furthermore, by generating diversified reputational insights, this approach enriches regulatory oversight, offering a multifaceted reference framework crucial for governance and system integrity.

c. Multimodal and Human-in-the-Loop Collaboration

With the rapid advancement of multimodal perception and generation technologies, multi-agent collaboration tasks are no longer limited to structured text processing. Instead, they increasingly extend to complex collaborative scenarios involving multimodal information such as images, audio, and tabular data [30, 31]. Therefore, to extend the scenario-generalizability of the proposed mechanism, future distributed multi-agent collaboration mechanisms must support multimodal input parsing, task reasoning, and cross-modal output validation to enhance their semantic consistency and task coverage in real-world applications. Furthermore, although the proposed mechanism emphasizes decentralization and autonomous execution, it remains essential to construct trustworthy interaction interfaces in human-in-the-loop collaborative tasks [32]. These interfaces should support task progress visualization, controllability feedback, and manual intervention access, introducing robust human oversight mechanisms for high-risk scenarios.

VI. Conclusion

This paper introduces a distributed multi-agent collaboration mechanism driven by smart contracts to address data monopolies and interoperability barriers. First, we proposed a validation method utilizing planning, execution, and validation smart contracts, which can reduce error propagation through the proof of the resources of the knowledge base and tool-using with state rollbacks and execution atomicity. Second, we established a DAOs-driven management method integrating an auction-based tool supplier matching model to dynamically update the contracts and optimize service quality within the ecosystem. Third, we constructed a hierarchical evaluation indicator system and the key technologies for future implementation and empirical benchmarking. Future work will implement the proposed mechanism as a working prototype and empirically deploy and evaluate its performance against existing multi-agent frameworks.

Acknowledgment

This work was supported by the Science and Technology Development Fund, Macao SAR (Nos. 0093/2023/RIA2, 0157/2024/RIA2, and 0145/2023/RIA3).

References

[1] X. Luo, A. Rechardt, G. Sun, K. K. Nejad, F. Yáñez, B. Yilmaz, K. Lee, A. O. Cohen, V. Borghesani, A. Pashkov et al., Large language models surpass human experts in predicting neuroscience results, *Nat. Hum. Behav.*, 2025, 9(2), 305–315.

[2] Z. Xi, W. Chen, X. Guo, W. He, Y. Ding, B. Hong, M. Zhang, J. Wang, S. Jin, E. Zhou et al., The rise and potential of large language model based agents: A survey, *Sci. China Inf. Sci.*, 2025, 68(2), 121101.

[3] N. J. Szymanski, B. Rendy, Y. Fei, R. E. Kumar, T. He, D. Milsted, M. J. McDermott, M. Gallant, E. D. Cubuk, A. Merchant et al., An autonomous laboratory for the accelerated synthesis of novel materials, *Nature*, 2023, 624(7990), 86–91.

[4] P. Pataranutaporn, V. Danry, J. Leong, P. Punpongsanon, D. Novy, P. Maes, and M. Sra, AI-generated characters for supporting personalized learning and well-being, *Nat. Mach. Intell.*, 2021, 3(12), 1013–1022.

[5] S. Feng, W. Ding, A. Liu, Z. Wang, W. Shi, Y. Wang, Z. Shen, X. Han, H. Lang, C.-Y. Lee et al., When one LLM drools, multi-LLM collaboration rules, arXiv preprint arXiv: 2502.04506, 2025.

[6] K. T. Tran, D. Dao, M. D. Nguyen, Q. V. Pham, B. O'Sullivan, and H. D. Nguyen, Multi-agent collaboration mechanisms: A survey of LLMs, arXiv preprint arXiv: 2501.06322, 2025.

[7] M. Cohen and A. Sundararajan, Self-regulation and innovation in the peer-to-peer sharing economy, *Univ. Chicago Law Rev. Online*, 2015, 82(1), 8.

[8] E. G. Weyl, A price theory of multi-sided platforms, *Am. Econ. Rev.*, 2010, 100(4), 1642–1672.

[9] J. Li, B. Li, X. Liu, R. Xu, J. Ma, and H. Yu, Breaking data silos: Cross-domain learning for multi-agent perception from independent private sources, in *Proc. 2024 IEEE International Conference on Robotics and Automation*, Yokohama, Japan, 2024, 18414–18420.

[10] M. Tang, F. Peng, and V. W. S. Wong, A blockchain-empowered incentive mechanism for cross-silo federated learning, *IEEE Trans. Mob. Comput.*, 2024, 23(10), 9240–9253.

[11] M. Zhang, F. Lin, L. Tian, R. Jia, Z. Zheng, and M. Li, Evolutionary medical data modeling and sharing via federated learning over sharded blockchain, *IEEE Internet Things J.*, 2024, 11(13), 24234–24246.

[12] L. Cao, Decentralized AI: Edge intelligence and smart blockchain, metaverse, Web3, and DeSci, *IEEE Intell. Syst.*, 2022, 37(3), 6–19.

[13] S. Hong, M. Zhuge, J. Chen, X. Zheng, Y. Cheng, C. Zhang, J. Wang, Z. Wang, S. K. S. Yau, Z. Lin et al., MetaGPT: Meta programming for a multi-agent collaborative framework, in *Proc. 12th International Conference on Learning Representations*, Vienna, Austria, 2024, 1–29.

[14] Y. Yu, Z. Yao, H. Li, Z. Deng, Y. Jiang, Y. Cao, Z. Chen, J. W. Suchow, Z. Cui, R. Liu et al., FINCON: A synthesized LLM multi-agent system with conceptual verbal reinforcement for enhanced financial decision making, in *Proc. 38th International Conference on Neural Information Processing Systems*, Vancouver, Canada, 2024, 4354.

[15] M. Cemri, M. Pan, S. Yang, L. A. Agrawal, B. Chopra, R. Tiwari, K. Keutzer, A. Parameswaran, D. Klein, K. Ramchandran et al., Why do multi-agent LLM systems fail? arXiv preprint arXiv: 2503.13657, 2025.

[16] E. Chang and L. Geng, SagaLLM: Context management, validation, and transaction guarantees for multi-agent LLM planning, *Proc. VLDB Endowment*, 2025, 18(12), 4874–4886.

[17] B. Pan, J. Lu, K. Wang, L. Zheng, Z. Wen, Y. Feng, M. Zhu, and W. Chen, AgentCoord: Visually exploring coordination strategy for LLM-based multi-agent collaboration, *Comput. Graph.*, 2025, 132, 104338.

[18] J. Li, R. Qin, S. Guan, J. Hou, and F.-Y. Wang, Blockchain intelligence: Intelligent blockchains for Web 3.0 and beyond, *IEEE Trans. Syst., Man, Cybern.: Syst.*, 2024, 54(11), 6633–6642.

[19] S. Wang, L. Ouyang, Y. Yuan, X. Ni, X. Han, and F.-Y. Wang, Blockchain-enabled smart contracts: Architecture, applications, and future trends, *IEEE Trans. Syst., Man, Cybern.: Syst.*, 2019, 49(11), 2266–2277.

[20] R. Qin, W. Ding, J. Li, S. Guan, G. Wang, Y. Ren, and Z. Qu, Web3-based decentralized autonomous organizations and operations: Architectures, models, and mechanisms, *IEEE Trans. Syst., Man, Cybern.: Syst.*, 2023, 53(4), 2073–2082.

[21] S. Hu, C. Lu, and J. Clune, Automated design of agentic systems, in *Proc. 13th International Conference on Learning Representations*, Singapore, Singapore, 2025, 1–34.

[22] J. Zhang, J. Xiang, Z. Yu, F. Teng, X.-H. Chen, J. Chen, M. Zhuge, X. Cheng, S. Hong, J. Wang et al., AFlow: Automating agentic workflow generation, in *Proc. 13th International Conference on Learning Representations*, Singapore, Singapore, 2025, 1–38.

[23] W. Ding, J. Li, R. Qin, S. Guan, and F.-Y. Wang, AI4S based on DeSci: Reference model and research issues, in *Proc. 2023 IEEE International Conference on Systems, Man, and Cybernetics*, Honolulu, HI, USA, 2023, 4784–4789.

[24] F.-Y. Wang, R. Qin, J. Li, X. Wang, H. Qi, X. Jia, and B. Hu, Federated management: Toward federated services and federated security in federated ecology, *IEEE Trans. Comput. Soc. Syst.*, 2021, 8(6), 1283–1290.

[25] D. H. Shih, B. Lin, and S.-Y. Huang, MoRVAM: A reverse Vickrey auction system for mobile commerce, *Expert Syst. Appl.*, 2007, 32(4), 1113–1123.

[26] W. Vickrey, Counterspeculation, auctions, and competitive sealed tenders, *J. Finance*, 1961, 16(1), 8–37.

[27] I. Chakraborty, F. Khalil, and J. Lawarree, Competitive procurement with ex post moral hazard, *RAND J. Econ.*, 2021, 52(1), 179–206.

[28] D. D. H. Shin, Blockchain: The emerging technology of digital trust, *Telemat. Inform.*, 2019, 45, 101278.

[29] X. Xue, S. Wang, L. Zhang, Z. Feng, and Y. Guo, Social learning evolution (SLE): Computational experiment-based modeling framework of social manufacturing, *IEEE Trans. Industr. Inform.*, 2019, 15(6), 3343–3355.

[30] F. Lin, T. Gao, D. Sun, Q. Ni, X. Ding, J. Wang, D. Gao, and F.-Y. Wang, Parallel medical devices and instruments: Integrating edge and cloud intelligence for smart treatment and health systems, *IEEE/CAA J. Autom. Sin.*, 2025, 12(4), 651–654.

[31] F. Lin, Y. Tian, Y. Wang, T. Zhang, X. Zhang, and F.-Y. Wang, AirVista: Empowering UAVs with 3D spatial reasoning abilities through a multimodal large language model agent, in *Proc. 2024 IEEE 27th International Conference on Intelligent Transportation Systems*, Edmonton, Canada, 2024, 476–481.

[32] W. Takerngsaksiri, J. Pasuksmit, P. Thongtanunam, C. Tantithamthavorn, R. Zhang, F. Jiang, J. Li, E. Cook, K. Chen, and M. Wu, Human-in-the-loop software development agents, in *Proc. 2025 IEEE/ACM 47th International Conference on Software Engineering: Software Engineering in Practice*, Ottawa, Canada, 2025, 342–352.

Sangtian Guan received the ME degree in intelligent technology from Macau University of Science and Technology, China, in 2023. He is currently pursuing the PhD degree at Department of Engineering Science, Faculty of Innovation Engineering, Macau University of Science and Technology, Macao, China. His research interests include DeEco, blockchain, DAO, and parallel management.

Fei Lin received the ME degree in intelligent technology from Macau University of Science and Technology, China, in 2023. He is currently pursuing the PhD degree at Department of Engineering Science, Faculty of Innovation Engineering, Macau University of Science and Technology, Macao, China. His research interests include parallel intelligence, large language model, and embodied agent.

Juanjuan Li received the MS degree in economics from Renmin University of China, China, in 2010, and the PhD degree in control science and engineering from Beijing Institute of Technology, Beijing, China, in 2021. She is currently an associate professor at State Key Laboratory of Multimodal Artificial Intelligence Systems, Institute of Automation, Chinese Academy of Sciences, Beijing, China. Her research interests include blockchain, DAO, and decision intelligence.

Xiaolong Liang received the BS and MS degrees in computer science and technology from Shandong University, Jinan, China, in 2011 and 2016, respectively, and the PhD degree in intelligent science and systems from Macau University of Science and Technology, Macao, China, in 2024. He is currently an assistant professor at State Key Laboratory of Multimodal Artificial Intelligence Systems, Institute of Automation, Chinese Academy of Sciences, China. His research interests include blockchain intelligence, DAOs (TAOs), intelligent transportation systems, and parallel governance.