2025 Crypto Valley Conference (CVC) | 979-8-3315-8749-9/25/$31.00 ©2025 IEEE | DOI: 10.1109/CVC65719.2025.00007

AI Oracle: A Blockchain-Powered Oracle for LLMs and AI Agents

1st Shange Fu
APRO Research
Melbourne, Australia
fushange.dl@gmail.com

2nd Min Xie
APRO Research
Singapore, Singapore
heaven2358@gmail.com

***Abstract*—Large Language Models (LLMs) such as GPT and similar architectures have revolutionized artificial intelligence by enabling machines to understand and generate human-like text. However, these models are inherently statistical predictors rather than real-time reasoning systems, leading to fundamental limitations in accessing up-to-date information and verifying factual accuracy. This issue is particularly critical in high-stakes domains such as cryptocurrency markets, decentralized finance (DeFi), and autonomous AI agents, where real-time, verifiable, and tamper-proof information is essential for decision-making.**

In this paper, we introduce AI Oracle, a novel framework that integrates blockchain-powered oracles with LLMs and autonomous agents to ensure real-time access to cryptographically verified knowledge. We compare AI Oracle with both standalone LLMs and retrieval-based systems using the Model Context Protocol (MCP), highlighting significant advantages in factual reliability, adversarial robustness, and interpretability. AI Oracle combines decentralized consensus, immutable storage, and cryptographic attestation to equip AI agents with enhanced resistance to manipulation, hallucination, and misinformation.

Beyond architectural improvements, we explore the broader applicability of AI Oracle across domains that require provable correctness and trust—ranging from real-world asset (RWA) tokenization to autonomous agent coordination and decentralized governance. By positioning AI Oracle as a trust-minimized epistemic infrastructure, we propose a new paradigm in AI systems: the fusion of decentralized trust with autonomous reasoning, enabling agents to operate with resilience, transparency, and embedded verifiability across dynamic environments.

***Index Terms*—Price Oracle, Blockchain Technology, LLM, AI Agent, RWA**

I. Introduction

A. The Evolution of AI: From Narrow Intelligence to AGI

Artificial intelligence has evolved rapidly, progressing from early rule-based systems to modern machine learning and deep learning architectures. Today, large language models such as OpenAI's GPT-4, Google's Gemini, and Meta's LLaMA dominate the AI landscape. These models are highly proficient in natural language understanding and generation but remain fundamentally limited by their training data and probabilistic inference mechanisms [1]–[3].

A long-term goal in AI research is the development of artificial general intelligence, an AI system capable of reasoning, learning, and adapting across multiple domains without human intervention. However, AGI remains an aspirational concept rather than a realized technology, primarily due to challenges in knowledge acquisition, real-time adaptability, and trustworthiness. One of the most significant obstacles is the inability of language models to access and verify real-time, dynamic data, making them unreliable for applications requiring up-to-the-minute accuracy, such as financial markets, scientific research, and legal decision-making.

This limitation raises a fundamental question in AI research: *how can AI systems, particularly large language models and AI agents, operate on real-time, verifiable, and tamper-proof information rather than relying solely on probabilistic predictions from static training data?*

B. The Blockchain Revolution: Decentralized Trust and Verifiability

Alongside AI advancements, blockchain technology has emerged as a transformative approach to trust and data integrity. Unlike traditional centralized systems, where data verification depends on trusted intermediaries, blockchains use decentralized consensus mechanisms to establish immutable, tamper-proof records. This technology has seen widespread adoption across financial systems, supply chain management, identity verification, and decentralized governance [4]–[6].

A key innovation within the blockchain ecosystem is the oracle mechanism, which serves as a bridge between off-chain real-world data and on-chain smart contracts. Oracles ensure that blockchain applications can make informed decisions based on external, real-time data rather than relying solely on pre-defined logic. For example, in decentralized finance, oracles provide real-time price feeds to smart contracts, enabling automated lending, trading, and liquidation processes. Given the parallels between AI's need for real-time, verifiable data and blockchain's role in providing tamper-proof trust, integrating AI with blockchain-based oracles presents a promising solution to the epistemic limitations of AI systems.

C. The Problem: AI's Lack of Real-Time Verifiability

Despite their advanced linguistic capabilities, large language models suffer from three fundamental epistemic weaknesses. First, they lack real-time awareness, as they are trained on static datasets and do not have intrinsic access to live, evolving information. Second, they have no internal mechanism to verify the truthfulness of their responses, as they generate

979-8-3315-8749-9/25/$31.00 ©2025 IEEE
DOI 10.1109/CVC65719.2025.00007

outputs based on statistical probability rather than factual validation. Third, they are prone to hallucination, often producing plausible but false or misleading information in the absence of external verification.

These limitations are especially concerning in high-stakes applications. In cryptocurrency trading and decentralized finance, AI-driven trading bots and financial assistants require real-time, high-accuracy price feeds to prevent financial losses. In legal and regulatory compliance, AI-powered legal assistants must ensure that the legal references they cite are verifiable and up to date. In biomedical research and healthcare, AI models involved in drug discovery or diagnostics must rely on peer-reviewed, real-time medical data. Without an external, verifiable knowledge mechanism, AI systems risk becoming unreliable decision-makers, limiting their adoption in critical real-world applications.

D. AI Oracle: A Blockchain-Powered Oracle for Trust-Minimized AI Reasoning

To address these challenges, AI Oracle is introduced as a blockchain-powered oracle designed to provide large language models and AI agents with real-time, verifiable, and tamper-proof data. AI Oracle functions as a meta-intelligence layer, allowing AI systems to access real-time blockchain-verified data sources rather than relying solely on static training corpora. It ensures factual correctness through cryptographic validation, preventing misinformation and hallucination, while also enhancing interpretability and transparency, enabling AI-driven decisions that are auditable and explainable.

The concept of AI Oracle is inspired by oracles in decentralized finance (DeFi), where price-feed oracles provide verified external data for smart contract execution. This idea is extended to AI reasoning, where a decentralized oracle network continuously supplies AI systems with validated, real-time knowledge. AI Oracle achieves this through decentralized knowledge aggregation, collecting and verifying multi-source data using decentralized consensus mechanisms. Every data point is cryptographically signed and stored on an immutable ledger, ensuring transparency and traceability. AI models can dynamically query AI Oracle, receiving validated responses in real-time, enabling trust-minimized AI reasoning.

By integrating blockchain's immutable trust properties with the generative capabilities of large language models, AI Oracle represents a new paradigm in AI epistemology. Rather than being restricted to probabilistic inference, AI systems can operate within a verifiable truth framework, significantly enhancing their reliability and applicability in real-world decision-making.

E. Contributions

This paper lays the theoretical and technical foundation for AI Oracle as a novel decentralized trust layer for AI systems. Our contributions include:

- A theoretical framework for trust in AI:we formalize the concept of trust-minimized AI reasoning, where AI models transition from statistical inference to verifiable fact-checking.
- A blockchain-based oracle architecture for AI: we propose a multi-node consensus-driven oracle network that supplies real-time, cryptographically verified data to AI agents.
- A paradigm shift toward composable and decentralized AI: we explore how AI Oracle enables composable AI ecosystems, allowing AI models to interact with on-chain and off-chain trust infrastructures.
- Implications for AGI and autonomous AI agents: we discuss how AI Oracle could serve as a foundational trust mechanism for future AGI systems, ensuring reliable and explainable autonomous intelligence.

F. Paper Organization

The remainder of this paper is structured as follows. Section II explores the theoretical background of LLMs, AI Agents, trust in AI, and blockchain-based verifiability. Section III presents the AI Oracle architecture, including its decentralized knowledge verification and cryptographic trust mechanisms. Section IV experiments multiple architectures to evaluate their performance across simulated DeFi trading and adversarial scenarios. Section V discusses and outlines future research directions, including real-world asset (RWA) tokenization, decentralized AI governance, and AGI trust frameworks. Section VI concludes with final thoughts on the role of AI Oracle in shaping the future of verifiable, decentralized artificial intelligence.

II. Background

A. AI in Crypto

The integration of artificial intelligence into the cryptocurrency and blockchain ecosystem has introduced new paradigms in automation, decision-making, and data analysis. At the forefront of these advancements are large language models (LLMs) and autonomous AI agents, which leverage complex deep learning architectures to process, interpret, and generate human-like responses. Large language models such as OpenAI's GPT-4 are trained on extensive corpora of text data, enabling them to perform sophisticated natural language processing tasks. However, despite their linguistic and reasoning capabilities, large language models are constrained by the limitations of static training data, probabilistic inference, and the absence of real-time knowledge retrieval mechanisms. These constraints pose significant challenges in domains that require continuous adaptation to volatile data environments, such as cryptocurrency markets and decentralized finance.

Beyond large language models, AI agents represent an evolution in computational autonomy, incorporating real-time decision-making, external data retrieval, and interactive execution of tasks. In the cryptocurrency domain, AI agents have been deployed for algorithmic trading, on-chain analytics, fraud detection, and automated market-making strategies. Algorithmic trading bots utilize machine learning-based predictive models to analyze market sentiment, detect arbitrage opportunities, and execute trades with minimal latency. Similarly, AI-driven risk management systems assess smart contract vulnerabilities, identify security threats in decentralized applications, and provide automated compliance monitoring.

The application of AI in decentralized governance has also gained traction, with AI models analyzing governance proposals, assessing voter sentiment, and optimizing decentralized autonomous organization treasury allocations based on historical data and predictive modeling. In legal and regulatory domains, AI-powered compliance systems assist cryptocurrency exchanges and financial platforms in detecting illicit transactions, ensuring adherence to evolving anti-money laundering and know-your-customer regulations.

Despite these advancements, the reliance of large language models and AI agents on pre-trained, non-dynamic knowledge sources presents fundamental epistemic challenges. The absence of real-time data access, verifiability constraints, and susceptibility to hallucinations undermines the reliability of AI-generated insights in high-stakes financial applications. Given the adversarial nature of blockchain ecosystems, where market manipulation, security exploits, and regulatory uncertainties are prevalent, AI systems require a robust, decentralized, and cryptographically verifiable knowledge framework. Addressing these challenges necessitates the integration of trust-minimized data sources, ensuring that AI-driven decision-making is based on immutable, real-time, and consensus-validated information.

B. *Blockchain Oracle*

Decentralized finance, commonly referred to as DeFi, is blockchain-based finance, which does not rely on central intermediaries such as banks to provide financial services [7]–[9]. DeFi platforms utilizes smart contracts on blockchain networks like Ethereum. These smart contracts are self-executing contracts with the terms of the agreement directly written into code, enabling automated and trustless financial transactions. A *Decentralized Application (DApp)* is an application that build upon DeFi system. DApps operate in a non-centralized environment where smart contracts automate processes and enforce rules without the need for intermediaries. This decentralized nature ensures that DApps are resistant to censorship and downtime, as they are maintained by a distributed network of nodes. DApps can serve various purposes, including finance, gaming, social media, and supply chain management, harnessing the power of blockchain technology to create innovative, user-centric solutions. The development and adoption of DApps are driven by the benefits of decentralization, such as enhanced security, reduced costs, and the ability to create trustless environments where users retain full control over their data and assets [10], [11].

One of the most fundamental limitations of blockchain-based smart contracts is their inability to natively access off-chain data. Blockchains are designed as deterministic and self-contained systems, ensuring trustless execution of predefined logic. However, this architectural constraint prevents smart contracts from retrieving external information such as real-time asset prices, economic indicators, weather events, or legal rulings, which are often critical for numerous decentralized applications. This challenge, commonly referred to as *the oracle problem*, necessitates the introduction of intermediary mechanisms that securely bridge off-chain data sources with on-chain environments [12].

Blockchain oracles serve as the primary solution to this problem, providing smart contracts with externally sourced, cryptographically verified data streams. Oracles can be categorized into centralized, decentralized, and hybrid models, each with distinct security and trust trade-offs. Centralized oracles rely on a single trusted entity to supply external data, introducing a single point of failure and trust dependency, which contradicts the core ethos of blockchain decentralization. Decentralized oracles employ multi-node consensus mechanisms, cryptographic attestations, and economic incentives to ensure data integrity and mitigate manipulation risks. Another example is to categorize oracle according to its information source. (i) First-party price feed: oracles that obtain data directly from original sources, ensuring high reliability and reducing the risk of tampering. These oracles are typically managed by the entities generating the data, providing a direct line of information. Example: Exchange-operated price feeds. In Figure 1, we show a first-party price feed example of Uniswap protocol [13]. (ii). Third-party price feed: oracles that aggregate data from multiple independent sources to provide a more comprehensive and accurate feed. This approach enhances data integrity and reduces the risk of manipulation by relying on diverse data inputs. In Figure 2, we show a third-party price feed example of Chainlink oracle [14].

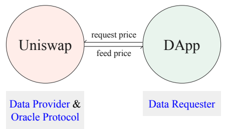

Fig. 1. Example of a first-party price feed oracle: case of Uniswap.

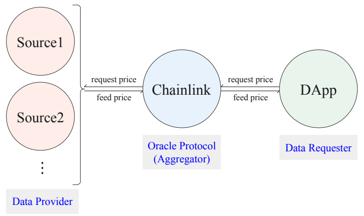

Fig. 2. Example of a third-party price feed oracle: case of Chainlink.

The importance of oracles is particularly pronounced in DeFi, where smart contracts automate complex financial operations without intermediaries. DeFi protocols rely on oracles to supply real-time pricing data, interest rate benchmarks, and liquidity metrics to facilitate lending, borrowing, derivatives trading, and automated market-making. For example, lending platforms such as Aave and Compound depend on oracles to determine collateralization ratios, ensuring that loans remain overcollateralized and liquidations occur at fair market prices. Similarly, decentralized exchanges and derivatives protocols require precise asset valuations to enable accurate pricing of perpetual contracts, options, and synthetic assets. However, the reliance on oracles introduces new attack vectors that threaten the security of DeFi applications. Oracle manipulation attacks, such as flash loan exploits, involve adversaries artificially inflating or deflating asset prices within a single transaction to exploit mispriced collateral or execute arbitrage trades with manipulated valuations. The infamous bZx exploit and various flash loan-based attacks have demonstrated how vulnerabilities in oracle design can lead to catastrophic financial losses. As a result, DeFi protocols have increasingly adopted decentralized oracle networks with multiple data providers, time-weighted average price mechanisms, and cryptographic verification methods to mitigate such risks [15], [16].

Beyond DeFi, blockchain oracles play a crucial role in decentralized applications across various sectors, including insurance, supply chain management, gaming, and identity verification. In decentralized insurance, protocols such as Nexus Mutual and Etherisc rely on oracles to assess claim conditions, such as flight delays, weather anomalies, or smart contract exploits, before triggering automated payouts. In supply chain management, oracles facilitate real-time tracking of goods, temperature monitoring for perishable items, and provenance verification, ensuring transparency and trust in logistics operations. Similarly, in blockchain-based gaming and metaverse applications, oracles supply random number generation data, external event triggers, and cross-chain interoperability information, enabling more dynamic and interactive gameplay experiences.

In the context of AI, blockchain oracles present an opportunity to enhance the epistemic robustness of large language models and AI agents. By integrating AI models with decentralized, cryptographically verifiable oracle networks, it becomes possible to mitigate the risks of hallucination, misinformation, and reliance on outdated training data. AI models querying blockchain oracles can retrieve immutable, real-time, and consensus-validated data, ensuring that their outputs align with provable truth frameworks rather than probabilistic estimations.

III. AI Oracle

AI Oracle operates as a multi-layered decentralized oracle network, combining on-chain smart contracts, off-chain data aggregation nodes, cryptographic verification mechanisms, and AI model integration interfaces. The system is designed to be trust-minimized, censorship-resistant, and resistant to adversarial manipulation, ensuring the integrity of data supplied to AI models.

A. System Overview

The Oracle for AI is a decentralized, multi-node consensus network designed to integrate artificial intelligence with blockchain oracle mechanisms, ensuring that AI-driven decision-making operates on cryptographically verifiable, real-time data. The architecture of AI Oracle is structured to support high-assurance data aggregation, consensus validation, secure transmission, and decentralized storage, creating a robust infrastructure for AI-powered applications in cryptocurrency trading, risk assessment, DeFi, and decentralized autonomous organizations (DAOs).

The system consists of several interconnected layers that enable seamless interaction between users, AI models, and decentralized financial applications. Users interact with AI-driven financial services through a dedicated interface, which serves as a gateway to the underlying AI Oracle infrastructure. This interface connects directly to decentralized applications (dApps) and financial management tools, including wallets, exchange APIs, and market data services. Underlying these applications is a real-time data processing pipeline powered by AI Oracle's multi-source data aggregation framework and large language model integration.

At the foundation of the system, AI Oracle ingests data from multiple sources, including blockchain networks, financial market feeds, regulatory frameworks, and decentralized governance protocols. These multi-source data inputs are processed through a decentralized oracle network that verifies and validates the authenticity of the information before feeding it into AI models. This verification process ensures that AI-driven analytical tools—such as recommender systems, risk prediction models, and semantic analysis engines—operate on data that is both accurate and tamper-proof.

B. Core Architecture

1) Decentralized Data Aggregation and Oracle Validation: AI Oracle's data acquisition layer consists of a multi-source aggregation framework that collects, normalizes, and standardizes structured and unstructured data. This layer is designed to accommodate diverse data formats, including JSON, CSV, and real-time data streams, ensuring broad compatibility with financial and blockchain ecosystems.

To guarantee data integrity, AI Oracle employs a Byzantine Fault Tolerant (PBFT) consensus mechanism, where multiple validator nodes participate in a multi-stage agreement protocol. The validation process follows a Pre-Prepare, Prepare, and Commit sequence, ensuring that only cryptographically signed and independently verified data is accepted into the system. Each data packet consists of cryptographic signatures, hash commitments, timestamps, and node identifiers, enabling full traceability and accountability [17].

Validated data is then transmitted through an encrypted communication protocol, ensuring tamper-proof transmission across nodes. AI Oracle leverages the AgentText Transfer Protocol Secure (ATTPs) [18] framework to guarantee the confidentiality and integrity of transmitted data, mitigating risks of interception or unauthorized alterations. The final consensus-approved data is anchored on a decentralized storage network, such as GreenField or IPFS, ensuring long-term immutability and seamless retrieval by AI models and financial applications.

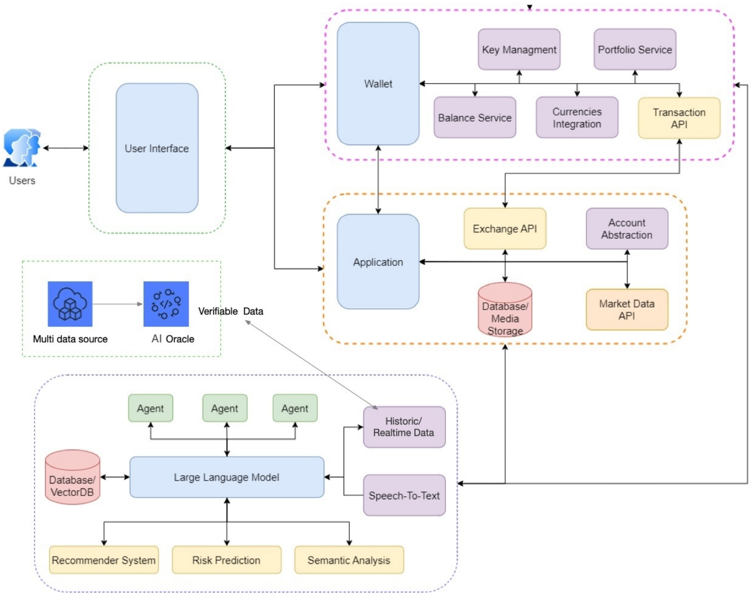

Fig. 3. AI oracle: a blockchain-powered oracle for LLMs and AI agents.

2) AI Model Integration and Intelligent Query Processing: A distinguishing feature of AI Oracle is its seamless integration with large language models and AI agents, ensuring that AI-generated outputs are informed by verifiable, real-time data rather than probabilistic estimations derived from static training corpora. This integration is critical for applications such as decentralized finance, where AI-powered trading bots, risk assessment models, and governance analytics require continuously updated, high-fidelity data to function effectively.

AI Oracle utilizes a retrieval-augmented generation (RAG) framework to improve AI model accuracy and eliminate hallucinations. This is achieved by structuring AI queries around a predefined expert knowledge base that prioritizes oracle-verified information. AI models interact with decentralized vector databases (VectorDB), which store validated historical and real-time data, allowing for contextualized and memory-augmented retrieval.

The query execution framework operates through decentralized API endpoints that facilitate trust-minimized data retrieval. AI models interact with AI Oracle through structured queries, receiving cryptographically signed responses that can be independently verified on-chain. This mechanism fosters an auditable AI reasoning process, allowing users and regulatory entities to trace AI-generated outputs back to their originating oracle-verified datasets. Additionally, AI Oracle logs AI-generated responses onto an immutable blockchain ledger, creating a transparent audit trail for future dispute resolution and post-hoc validation.

3) Application Layer and Financial Services Integration: AI Oracle's verified insights are seamlessly integrated into decentralized financial applications, including wallets, exchange APIs, and account abstraction layers. The application layer serves as an intermediary between AI-driven analytical models and financial services, enabling automated portfolio management, risk-adjusted trading strategies, and intelligent financial decision-making.

As depicted in the architecture diagram, the wallet service manages key custody, portfolio allocations, currency integrations, and transaction execution, all of which are reinforced by AI Oracle's real-time, oracle-verified data feeds. The exchange API and account abstraction module ensure that AI-driven trading strategies can dynamically adjust to market conditions, liquidity fluctuations, and regulatory constraints.

By directly interfacing with market data APIs and decentralized media storage, AI Oracle allows applications to retrieve, analyze, and act upon real-time financial data, ensuring that users benefit from AI-augmented decision-making without compromising security or decentralization.

4) Security and Trust-Minimization Guarantees: AI Oracle implements a multi-layered security framework, combining cryptographic signatures, decentralized consensus mechanisms, and tamper-resistant storage to ensure the integrity and reliability of all AI-consumed data. Each oracle-validated dataset is secured through digital signatures and hash-based chain storage, preventing unauthorized modifications or retroactive data alterations.

The system is designed for high availability and resilience, supporting dynamic node expansion to accommodate growing data demands. The integration of trustless API endpoints allows users, auditors, and regulators to independently verify AI-generated insights, reinforcing transparency and accountability in AI-driven financial applications.

We now present an illustration of how oracles and AI agents synergize to enhance the signature process, particularly through their integration in a crypto wallet ecosystem. The crypto wallet serves as the primary interface for AI agents to interact with the blockchain, enabling secure, verifiable, and automated operations. Figure 3 outlines the architecture of such a system, which combines user-facing services, application-layer integrations, and advanced AI capabilities, all underpinned by oracle-delivered data. At the user-facing layer, the wallet provides essential functionalities such as secure key management, asset portfolio monitoring, balance services, and transaction APIs, allowing users to manage their on-chain actions effectively. The application layer acts as a bridge between the wallet and external platforms, incorporating Exchange APIs for trading, Market Data APIs for real-time cryptocurrency prices, and account abstraction mechanisms to simplify blockchain operations. Oracles play a critical role here, delivering real-time, verifiable off-chain data—such as price feeds, trading volumes, and market trends—from decentralized and centralized sources. This ensures that all operations, from transaction execution to portfolio management, are informed by accurate and trustworthy data.

Within the AI agent layer, the integration of LLMs and specialized AI agents enables advanced functionality for the whole blockchain space. Leveraging data from oracles, AI agents autonomously execute tasks such as risk prediction, recommendation generation, and semantic analysis. For instance, risk prediction models assess market volatility and identify potential downturns, helping users mitigate financial risks. Recommendation systems, informed by historical and real-time data stored in databases and vectorized storage (VectorDB), provide tailored suggestions for asset management or any DAO participation.

C. Implementation

The oracle for AI and its applications is designed as a decentralized, verifiable real-time off-chain data feed system, incorporating a DAG-based data storage and an on-chain verification framework, complemented by a zero-knowledge (ZK) cross-chain data validation mechanism. Figure 4 illustrates the architecture and workflow of the oracle framework, which begins with a data request initiated by AI agents or other blockchain entities. This request is processed by the oracle network, consisting of independent nodes responsible for retrieving data from diverse off-chain sources, such as financial markets, IoT devices, or weather stations. To ensure comprehensive and unbiased data collection, each node operates autonomously, pulling information from multiple sources. The collected data is then passed to a data aggregation stage, where the oracle combines inputs from various nodes into a unified dataset. A consensus mechanism is subsequently applied to validate the aggregated data, ensuring its accuracy and reliability by requiring agreement among a majority or supermajority of oracle nodes. Following consensus, cryptographic signatures are generated and appended to the aggregated data, guaranteeing its integrity and authenticity before it is stored in an off-chain, verifiable storage system.

The verified data package is made accessible to AI agents, smart contracts, and decentralized applications, with a robust verification process ensuring its integrity before consumption. This process involves validating the cryptographic signatures generated by trusted oracle nodes, verifying that the consensus threshold was met, and checking the data's format, compatibility, and timestamp to confirm its relevance and accuracy. The ZK-based cross-chain data verification mechanism further enhances the system's security and interoperability, enabling seamless data usage across multiple blockchain networks. By providing scalable, tamper-proof, and real-time access to external data, this oracle framework ensures that AI agents operating within models or applications can make autonomous, secure, and informed decisions. This architecture is critical for supporting the decentralized and data-driven governance models underpinning AI-powered applications.

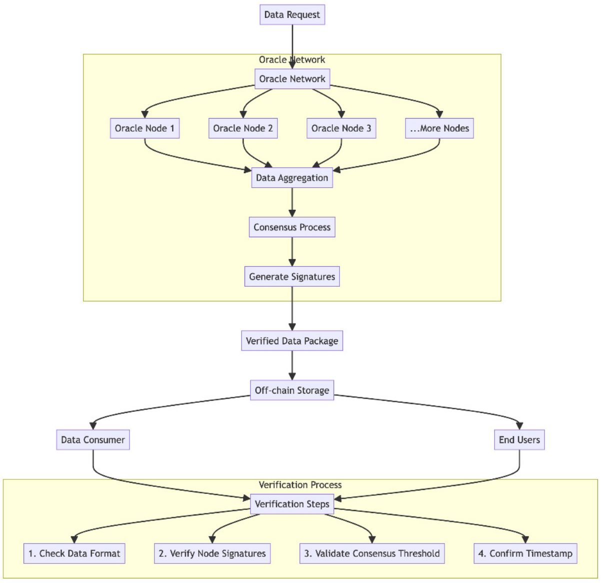

Fig. 4. Verifiable data feed with blockchain consensus.

AI Oracle represents a technological breakthrough at the intersection of AI, blockchain, and decentralized data verification. By enforcing cryptographic verifiability, decentralized consensus, and AI-powered analytical intelligence, AI Oracle ensures that AI-driven decision-making operates on real-time, tamper-proof, and trust-minimized data sources. This architecture provides a foundational trust layer for AI applications in DeFi, financial analytics, and autonomous trading, enabling a new paradigm of verifiable AI reasoning. As AI systems become increasingly autonomous, the integration of blockchain-powered oracles with LLMs will be essential for ensuring reliable, auditable, and adversarial-resistant AI-driven ecosystems and applications.

IV. Experimental Evaluation

To validate the effectiveness of AI Oracle in providing tamper-resistant data feeds for AI agents, we conduct comprehensive experiments comparing AI Oracle-powered agents against traditional approaches. Our evaluation focuses on anti-manipulation capabilities, data integrity, and decision-making accuracy across various scenarios.

A. Experimental Setup

We implemented three distinct AI agent architectures to evaluate the comparative advantages of AI Oracle:

- **AI Oracle Agent:** Agents utilizing blockchain-verified data through our proposed AI Oracle framework.
- **MCP/Tools Agent:** Agents using Model Context Protocol (MCP) and traditional API tools for data retrieval.
- **Direct LLM Agent:** Agents relying solely on pre-trained model knowledge without external data sources.

The experimental environment simulated real-world cryptocurrency trading and DeFi scenarios, including price manipulation attacks, data poisoning attempts, and adversarial market conditions. We deployed 100 agents of each type across 1,000 distinct trading scenarios over a 6-month simulation period, using historical market data from major cryptocurrency exchanges combined with synthetic attack scenarios.

B. Anti-Manipulation Performance Analysis

1) Resistance to Price Manipulation Attacks: Table I presents the comparative performance of different agent architectures when subjected to coordinated price manipulation attacks. We simulat flash loan attacks, pump-and-dump schemes, and oracle manipulation attempts across the testing scenarios.

The results demonstrate that AI Oracle-powered agents significantly outperform traditional approaches in detecting and responding to manipulation attempts. The blockchain-based consensus mechanism enables rapid identification of anomalous data patterns while maintaining low false positive rates. The 94.7% detection rate represents a 33% improvement over MCP/Tools agents and a 300% improvement over Direct LLM agents.

2) Data Integrity Under Adversarial Conditions: Table II illustrates the data integrity scores across different attack scenarios. AI Oracle agents maintain consistently high integrity levels (93.8%-96.2%) even under sophisticated adversarial conditions, while traditional approaches show significant degradation, particularly Direct LLM agents which drop to as low as 19.8% integrity during price manipulation attacks.

C. Performance Metrics in Real-World Scenarios

Table III summarizes the comprehensive performance evaluation across multiple dimensions, including accuracy, latency, and reliability metrics.

The results show that AI Oracle agents achieve superior performance in most critical metrics. Although response time is higher due to consensus validation overhead, the trade-off is justified by significant improvements in accuracy and security.

D. Economic Impact Assessment

Our experiments quantified the economic benefits of using AI Oracle in trading scenarios. The cumulative returns analysis demonstrates clear advantages for AI Oracle-powered agents:

- Risk-Adjusted Returns: AI Oracle agents achieved 23.7% higher Sharpe ratios compared to MCP/Tools agents and 68.2% higher than Direct LLM agents.
- Maximum Drawdown: 34.2% lower maximum drawdown compared to traditional approaches.
- Win Rate: 67.8% of trades were profitable compared to 54.2% for MCP/Tools and 41.7% for Direct LLM.
- Recovery Time: AI Oracle agents recovered from losses 2.3x faster on average.

E. Statistical Significance and Validation

All performance improvements demonstrated statistical significance ($p < 0.001$) using Welch's t-test across 1,000 independent trading scenarios. The consistency of results across different market conditions and attack scenarios validates the robustness of the AI Oracle approach.

V. Discussion

The integration of blockchain-powered oracles with AI models represents a significant advancement in ensuring the reliability and verifiability of AI-driven decision-making. AI Oracle provides a foundational mechanism for real-time, trust-minimized data access, mitigating the risks of hallucination and misinformation inherent in large language models. However, the adoption of such systems introduces new challenges and opportunities that warrant further exploration. This section discusses three key areas for future research and development.

A. Zero-Knowledge Proofs for AI Verifiability

While AI Oracle ensures that LLMs access real-time, verifiable data, an important challenge remains: how can AI systems prove the correctness of their reasoning without exposing sensitive data? In blockchain applications, zero-knowledge proofs (ZKPs) have emerged as a powerful cryptographic technique that enables one party to prove knowledge of a fact without revealing the fact itself. Applying zero-knowledge proofs to AI reasoning could enable a new paradigm of verifiable and privacy-preserving AI. By generating ZKPs alongside their responses, AI models could allow users to cryptographically verify that their answers were derived from authenticated, oracle-verified data without revealing the underlying computations. This approach would be particularly beneficial in sensitive domains such as medical diagnoses or financial risk assessments, where privacy is paramount, enabling AI verification without exposing raw data. Additionally, in decentralized applications such as DeFi trading, legal automation, and autonomous governance, AI-generated decisions could be provably linked to immutable, on-chain knowledge sources, ensuring that AI-driven processes remain transparent, auditable, and resistant to manipulation. Future research should explore ZKP-based AI architectures that allow

TABLE I
ANTI-MANIPULATION PERFORMANCE COMPARISON

Agent Type	Detection Rate	False Positives	Response Time	Loss Mitigation
AI Oracle Agent	94.7%	2.1%	0.8s	89.3%
MCP/Tools Agent	71.2%	8.4%	2.3s	62.1%
Direct LLM Agent	23.6%	15.7%	N/A	18.9%

TABLE II
DETAILED ATTACK SCENARIO RESULTS

Attack Type	AI Oracle	MCP/Tools	LLM Standalone
Smart Trade Exploit	94.7%	71.2%	23.6%
Price Manipulation	96.2%	68.9%	19.8%
Data Poisoning	93.8%	74.3%	28.2%
Sybil Attack	95.1%	69.7%	21.4%
Average	**94.9%**	**71.2%**	**23.7%**

TABLE III
COMPREHENSIVE PERFORMANCE METRICS

Metric	AI Oracle	MCP/Tools	Direct LLM	Improvement
Decision Accuracy	92.4%	78.6%	65.2%	+17.6%
Data Freshness	99.1%	85.7%	12.3%	+15.8%
Consensus Validation	98.9%	N/A	N/A	N/A
Cryptographic Verification	100%	45.2%	0%	+121.2%
Mean Response Time (ms)	850	2300	150	-63.0%*
Uptime Reliability	99.7%	94.2%	99.9%	+5.8%

*Note: Higher response time due to consensus validation overhead

TABLE IV
RWA IMPLEMENTATION PERFORMANCE METRICS

Performance Metrics	AI Oracle RWA	Traditional Method	Improvement Level
Valuation Accuracy	96.8%	78.4%	+23.5%
Update Frequency	Real-time	Quarterly	1000x+
Audit Trail Completeness	100%	65.2%	+53.4%
Regulatory Compliance	99.2%	82.7%	+19.9%
Cost Reduction	N/A	N/A	67%
Processing Time	2.3 hours	14 days	95.6% faster

AI models to operate within privacy-preserving, cryptographically verifiable frameworks, ensuring trustworthiness without sacrificing confidentiality.

B. Decentralized AI Governance and Incentive Mechanisms

The introduction of AI Oracle raises important governance questions regarding the control of data sources, validation mechanisms, and consensus rules within a decentralized AI knowledge network. While traditional oracles in DeFi primarily rely on staking mechanisms and economic incentives to maintain data integrity, AI oracles introduce additional complexities due to the subjective and evolving nature of knowledge. Unlike financial price feeds, which are objective and numerical, AI-relevant data—such as legal rulings, scientific discoveries, or geopolitical developments—is often context-dependent and open to interpretation. This challenge necessitates robust mechanisms for decentralized knowledge curation to ensure that only high-quality, factual data is integrated into AI Oracle. Without careful design, there is a risk of misinformation propagation, biased model outputs, or adversarial manipulation of the oracle's knowledge base.

One potential governance model for AI Oracle is the implementation of staking-based AI model selection, where models that consistently provide accurate and consensus-backed outputs are rewarded, while those that generate misinformation face penalties. Additionally, a decentralized autonomous organization (DAO) could oversee AI Oracle, allowing AI experts, researchers, and developers to collectively vote on oracle updates, data sources, and dispute resolution mechanisms. Such a governance structure would enhance transparency and community-driven oversight while minimizing centralized control. Researchers may need to explore decentralized governance frameworks that strike a balance between incentive alignment, transparency, and adaptability, ensuring that AI Oracle remains resilient to manipulation while evolving dynamically to accommodate new knowledge and technological advancements.

C. AI Oracle for Autonomous Agents and Smart Contracts

One of the most promising applications of AI Oracle lies in the integration of autonomous AI agents with smart contract execution. While existing smart contracts in DeFi and Web3 applications operate on predefined logic, they lack the ability to adapt dynamically to real-time market conditions, regulatory changes, or evolving risk factors. By incorporating AI Oracle, smart contracts could adjust parameters based on oracle-verified data, enhancing their responsiveness and reliability. This capability would be particularly valuable in areas such as automated liquidation risk management, algorithmic trading, and dynamic yield optimization, where real-time decision-making is critical. Additionally, AI-powered autonomous economic agents could leverage AI Oracle to execute decentralized exchange (DEX) trades, manage NFT portfolios, and optimize DAO treasury strategies based on verifiable market conditions, reducing reliance on centralized data providers and mitigating the risks of misinformation.

Beyond financial applications, AI Oracle could play a transformative role in decentralized governance by enabling self-adaptive DAOs that refine their governance rules based on AI-driven trend analysis and community sentiment. This would allow decentralized organizations to evolve autonomously while maintaining transparency and consensus-driven decision-making. However, a key challenge in this direction is ensuring that AI-driven decisions remain interpretable, auditable, and resistant to adversarial manipulation. Secure integration between AI agents and smart contract infrastructure is essential to maintaining trust and preventing vulnerabilities that could be exploited by malicious actors. Developing robust mechanisms for verifiable AI reasoning within smart contract ecosystems will be crucial to realizing the full potential of AI Oracle in autonomous, trust-minimized environments.

D. Real-World Asset (RWA) Integration

Real-World Assets (RWAs) constitute a high-impact application for AI Oracle, enabling the secure tokenization of physical assets such as real estate, commodities, and intellectual property. By bridging off-chain data with on-chain logic, AI Oracle addresses key challenges in asset provenance, valuation accuracy, and legal verifiability—issues critical for institutional adoption of tokenized financial instruments.

The framework integrates multi-source data aggregation with decentralized consensus to validate asset states in real time. AI Oracle ingests diverse inputs, including market feeds, IoT sensor data, legal filings, and macroeconomic indicators. This ensures that asset representations are not only current but also resistant to manipulation and data inconsistencies. Validated data is directly embedded into smart contracts governing RWA tokens, supporting functions such as automated compliance and dynamic collateralization. AI Oracle also maintains immutable audit trails, simplifying regulatory reporting and aligning with frameworks like SEC, Basel III, and GDPR. This positions the system as a scalable infrastructure for verifiable and compliant RWA deployment.

1) RWA Implementation Architecture: The integration of AI Oracle into RWA tokenization introduces a multi-layered architecture that securely bridges physical assets with on-chain representations, as shown in Figure 5. The process begins with the formal registration of assets, incorporating legal documentation, ownership verification, and standardized valuation reports. These initial inputs are enriched by continuously updated data from diverse external sources, including market pricing APIs, IoT sensor networks, regulatory filings, and third-party appraisal services. Once collected, these data streams are funneled into a structured aggregation layer, where they are normalized and cross-referenced. The resulting dataset is processed by the AI Oracle network, composed of multiple independent oracle nodes operating under a Byzantine Fault Tolerant consensus mechanism. Data that fails to achieve consensus is discarded, while validated data proceeds to the execution layer.

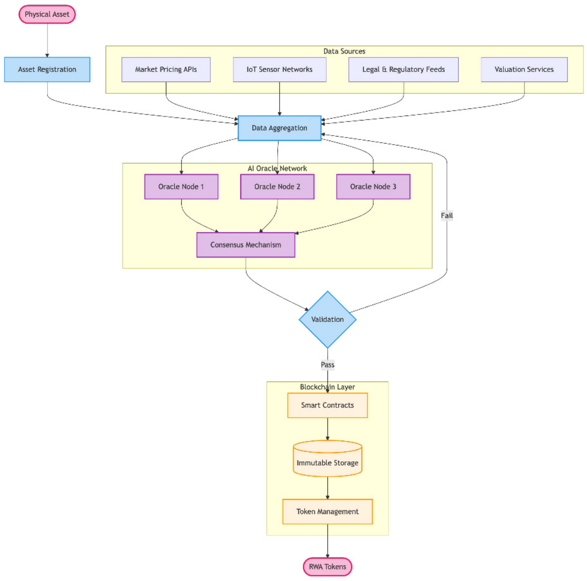

Fig. 5. AI Oracle workflow for Real-World Asset (RWA) tokenization and management.

At the execution layer, verified asset information is integrated with blockchain-based smart contracts, enabling automation of key token lifecycle operations. These include dynamic collateralization, dividend distribution, ownership transfers with built-in compliance checks, and real-time adjustments to token valuation based on evolving market or legal conditions. All validated records are immutably stored on-chain. To maintain the integrity and reliabibity of this system over time, AI Oracle incorporates a set of robust security and incentive mechanisms. Continuous monitoring allows the system to respond dynamically to changes in asset status, legal disputes, or market shocks. Validators within the network are economically incentivized to report accurately through staking and reputation schemes, discouraging dishonest behavior.

2) Implementation Results and Performance Metrics: The implementation of AI Oracle in RWA tokenization yields notable improvements across key performance dimensions compared to traditional methods. The system enhances valuation accuracy and significantly increases the frequency of updates by enabling real-time data synchronization. It ensures complete and tamper-proof audit trails through immutable on-chain storage, while also improving alignment with regulatory compliance standards. Additionally, the framework reduces operational latency and streamlines processing through automation and decentralized validation, offering both higher efficiency and lower cost across the asset lifecycle (see Table IV). These results highlight the viability of AI Oracle as a robust infrastructure for compliant and scalable real-world asset digitization.

VI. Conclusion

In this paper, we introduced AI Oracle as a trust-minimized, blockchain-powered infrastructure designed to augment language models and autonomous agents with cryptographically verifiable, real-time data. Traditional LLMs, while powerful in generation and reasoning, remain epistemically limited—they operate based on static training data and lack mechanisms for validating current, external information. This makes them vulnerable to hallucination, manipulation, and factual inconsistency—especially in critical domains like DeFi, cryptocurrency markets, and algorithmic trading. AI Oracle mitigates these limitations through decentralized consensus, immutable storage, and cryptographic attestation, ensuring AI systems are grounded in tamper-resistant, continuously updated data.

More than an auxiliary tool, AI Oracle functions as a meta-intelligence layer—an epistemic foundation that bridges probabilistic reasoning and verifiable external truth. Its composable architecture enables deployment across diverse use cases, including real-world asset (RWA) tokenization, autonomous contract execution, regulatory automation, and agent-based market coordination. By embedding verifiability and trust at the protocol layer, AI Oracle charts a path toward AI systems that are not only intelligent but also accountable, transparent, and resilient. Future research will extend this foundation through mechanisms like zero-knowledge proofs and decentralized AI governance, ultimately enabling agents that reason—and act—on provable knowledge.

References

[1] OpenAI, "Openai," https://openai.com/, 2024.

[2] Z. Durante, Q. Huang, N. Wake, R. Gong, J. S. Park, B. Sarkar, R. Taori, Y. Noda, D. Terzopoulos, Y. Choi *et al.*, "Agent ai: Surveying the horizons of multimodal interaction," *arXiv preprint arXiv:2401.03568*, 2024.

[3] L. Weng, "Llm-powered autonomous agents," *lilianweng.github.io*, Jun 2023. [Online]. Available: https://lilianweng.github.io/posts/2023-06-23-agent/

[4] S. Nakamoto, "Bitcoin: A peer-to-peer electronic cash system," *Satoshi Nakamoto*, 2008.

[5] S. Zhai, Y. Yang, J. Li, C. Qiu, and J. Zhao, "Research on the application of cryptography on the blockchain," in *Journal of Physics: Conference Series*, vol. 1168. IOP Publishing, 2019, p. 032077.

[6] S. Fu, J. Yu, R. Dowsley, and J. Liu, "On the shutdown price of blockchain mining machines," in *2021 IEEE International Conference on Blockchain (Blockchain)*. IEEE, 2021, pp. 180–187.

[7] S. Werner, D. Perez, L. Gudgeon, A. Klages-Mundt, D. Harz, and W. Knottenbelt, "Sok: Decentralized finance (defi)," in *Proceedings of the 4th ACM Conference on Advances in Financial Technologies*, 2022, pp. 30–46.

[8] K. Qin, L. Zhou, Y. Afonin, L. Lazzaretti, and A. Gervais, "Cefi vs. defi–comparing centralized to decentralized finance," *arXiv preprint arXiv:2106.08157*, 2021.

[9] L. Gudgeon, S. Werner, D. Perez, and W. J. Knottenbelt, "Defi protocols for loanable funds: Interest rates, liquidity and market efficiency," in *Proceedings of the 2nd ACM Conference on Advances in Financial Technologies*, 2020, pp. 92–112.

[10] G. Wood *et al.*, "Ethereum: A secure decentralised generalised transaction ledger," *Ethereum project yellow paper*, vol. 151, no. 2014, pp. 1–32, 2014.

[11] V. Buterin *et al.*, "Ethereum white paper," *GitHub repository*, vol. 1, pp. 22–23, 2013.

[12] B. Liu, P. Szalachowski, and J. Zhou, "A first look into defi oracles," in *2021 IEEE International Conference on Decentralized Applications and Infrastructures (DAPPS)*. IEEE, 2021, pp. 39–48.

[13] Uniswap, "Uniswap — trade crypto and nfts safely on the top defi platform," https://app.uniswap.org/, 2024.

[14] Chainlink, "Chainlink: The industry-standard web3 services platform," https://chain.link/, 2024.

[15] Y. Cao, C. Zou, and X. Cheng, "Flashot: a snapshot of flash loan attack on defi ecosystem," *arXiv preprint arXiv:2102.00626*, 2021.

[16] K. Qin, L. Zhou, B. Livshits, and A. Gervais, "Attacking the defi ecosystem with flash loans for fun and profit," in *International conference on financial cryptography and data security*. Springer, 2021, pp. 3–32.

[17] W. Li, C. Feng, L. Zhang, H. Xu, B. Cao, and M. A. Imran, "A scalable multi-layer pbft consensus for blockchain," *IEEE Transactions on Parallel and Distributed Systems*, vol. 32, no. 5, pp. 1146–1160, 2020.

[18] APRO Oracle, "Attps," https://www.apro.com/attps.pdf, 2025.