

Proof-of-Artificial-Intelligence-Work Protocol - Superlitepaper

Abstract

We propose Proof-of-Artificial-Intelligence-Work (PoAIW) as a novel consensus protocol that leverages artificial intelligence (AI) to secure and enable decentralized systems. Unlike traditional consensus mechanisms like Proof-of-Work (PoW)¹ or Proof-of-Stake (PoS)² but resembling Proof-of-Useful-Work (PoUW)³, PoAIW is the first consensus algorithm centered on the computational work performed by on-chain AI models. This superlitepaper outlines the concept, the technical foundation, and potential use cases for a fully autonomous Web3-native AI protocol.

Protocol Concept

The PoAIW protocol facilitates competition among on-chain AI models running as smart contracts and interacting under predefined protocol rules to achieve consensus and make progress in a decentralized, trustless, and round-based manner. Each round of consensus poses a protocol-defined challenge. On-chain AI models—hereafter called "mAIers"—compete by generating responses to the challenge. The protocol's ranking system evaluates these responses, and a set number of the best-performing mAIers are rewarded. The reward system encourages high-quality, efficient AI output and paves the way for intelligent computational results and decentralized system consensus.

Core Components

1. **Trustless AI Models:** AI models are required to run in a trustless manner for verifiable contributions and protocol adherence. Often, this will be achieved by running the AI entirely on-chain as smart contracts⁴. AI models may be used for protocol functionality (e.g. challenge generation, response evaluation, system monitoring, maintenance) and as mAIers which participate in the protocol (similar to traditional mining nodes).
2. **Rounds and Challenges:** Each round begins with a challenge generated according to the specific PoAIW implementation. Often, challenges will be AI-generated to ensure an autonomous protocol but certain use cases will require inputs from human actors or other IT systems. Depending on the use case and goals, the challenges will vary from answering questions to completing tasks or generating creative content. mAIers compete to deliver the best result based on these challenges.
3. **Ranking System:** Responses from mAIers are evaluated according to pre-defined protocol rules. Often, this will be achieved autonomously by a transparent on-chain system, which ranks them, e.g. based on response quality, relevance, generation

¹ <https://bitcoin.org/bitcoin.pdf> & https://en.wikipedia.org/wiki/Proof_of_work

² <https://decred.org/research/king2012.pdf>

³ https://wiki.internetcomputer.org/wiki/Proof_of_Useful_Work

⁴ For an example, see ICP's scalable canister architecture that enables resource-intense AI models, such as Large Language Models (LLMs), to execute computations securely, e.g. https://github.com/onicai/llama_cpp_canister

efficiency, or other set criteria. Protocol designers may opt for manual or hybrid evaluation processes as well. The protocol rules comprise round termination criteria, e.g. based on time or the number of response submissions, and how a round's winner is declared, e.g. the highest-ranked response or a set of best-performing AI models wins.

4. **Reward Mechanism:** Each winning mAIer mints a reward in accordance with the protocol's reward rules and schedule, similar to block rewards in traditional consensus mechanisms. Rewards can be paid in an existing cryptocurrency like BTC, ETH, or ICP (sourced from the competition pool) and in competition-specific tokens that are minted as part of the protocol.

Extendable Use Cases

PoAIW as a protocol is designed to be extendable to various AI-driven applications and competitions. Potential use cases include:

1. **Data Collection and Labeling:** The protocol could be used to collect high-quality training data, with AI models competing to label or generate synthetic data effectively.
2. **Federated AI Training:** New AI models may be created in a decentralized manner with a PoAIW implementation that leverages contributions in terms of data and computation by protocol participants.
3. **AI Model Selection⁵:** A competition where AI models compete to determine which one performs best on a specific task that streamlines model selection for developers and organizations.
4. **DeFi AI Trading Competitions:** AI models compete to generate the best trading strategies. The winning model mints a reward, creating a decentralized and autonomous trading competition⁶.
5. **Additional Ideas:** The protocol could be adapted for tasks in creative content generation, optimization problems, or any scenario where AI competition leads to high-quality outcomes.

Conclusion

With the Proof-of-Artificial-Intelligence-Work Protocol, we propose a novel approach to decentralized consensus and AI-driven competition. By leveraging on-chain AI models and tokenomics, PoAIW opens new possibilities for autonomous and intelligent blockchain protocols. The protocol's extendability ensures that it can serve a broad range of future use cases. A litpaper and a whitepaper as well as an open-source reference implementation will follow.

⁵ <https://github.com/IDEA-FinAI/LLM-as-a-Judge>

⁶ Also see discussions like [Decentralized Trading Competitions](#)