Quatern.ai

Unleashing the Power of Collective AI on Cardano.

Whitepaper v1.0

— February 2024 —

Abstract

This whitepaper presents *Quatern.ai*, an innovative platform built at the intersection of blockchain technology and Artificial Intelligence (AI) to revolutionize the way AI is developed, deployed, and utilized. In the face of the centralization issues plaguing current AI model management—such as concentration of power, lack of data privacy, and opaqueness of operations—*Quatern.ai* advocates for a decentralized and collaborative approach. By proposing a smart-contract-governed platform built on the Cardano blockchain, it not only puts forward a model of fairness and transparency but also underscores the superiority of collective AI models over individual counterparts. The essence of *Quatern.ai* lies in its conviction that collaboration among diverse AI models, facilitated through blockchain, can lead to more robust, accurate, and powerful AI solutions. The *$QAI* token is central to this ecosystem, rewarding contributions, enabling access to AI solutions, and granting the community governance power. Through this framework, *Quatern.ai* sets out to democratize AI, ensuring that its benefits are accessible to all, while enhancing the potency of AI through collective intelligence.

Contents

1 Introduction

1.1 Current Developments in AI

Propelled by advances in Machine Learning (ML), Artificial Intelligence (AI) has seen unprecedented growth in recent years and seems to be rapidly approaching the perhaps ultimate goal of attaining human-level intelligence. Moreover, the field has been quickly evolving from theoretical concepts to practical, impactful applications. Advancements in neural network architectures, and large-scale data processing have paved the way for AI to revolutionize a broad range of industries, from healthcare and agriculture to finance and logistics. In particular, latest breakthroughs in Large Language Models (LLMs) have made it evident that AI has arrived in the real world and is about the transform every aspect of the way we live, work, and communicate.

Like many experts in the field [Christian, 2020], we believe that we are currently entering a (possibly short) window of time in which the rules of the game for AI can still be set and controlled by us. A significant issue with the current method of deploying new AI models is that these models are typically managed, controlled, trained, etc., in a *centralized* manner. This centralized approach to AI raises several potential problems:

- **Concentration of Power:** Centralized AI model management can lead to a concentration of power in the hands of a few entities, stifling innovation and competition.

- **Data Privacy and Security Risks:** Centralization may pose risks to data privacy and security, as large, centrally controlled datasets are often required for training.

- **Single Point of Failure:** A central controller is a single point of failure, making AI systems more vulnerable to outages, data breaches, or malicious attacks.

- **Lack of Transparency and Accountability:** Centralized management often lacks transparency and accountability, hindering users' understanding of decision-making processes and the data that has been involved in training the model.

3

1.2 Decentralized & Collective AI

In response to the above listed challenges posed by centralized AI model management, we propose a shift towards a decentralized and collaborative approach, leveraging the strengths of blockchain technology. The benefits of such a paradigm shift are twofold:

- **Overcoming Centralization.** Decentralizing AI models via blockchain offers a robust solution to the problems inherent in centralized systems. It dilutes the concentration of power by distributing control across a network, thereby promoting innovation and competition. Additionally, a decentralized system reduces the risk of having a single point of failure, ensuring greater system stability and reliability.

- **Harnessing Collective Intelligence.** The advantages of decentralization in the context of AI actually go far beyond solving centralization issues. Such a distributed setup additionally enables the amalgamation of diverse data sources, model architectures, and training methodologies. By facilitating contributions from a wide array of participants, one can harness collective intelligence to form powerful ensemble models [Breiman, 1996]. These essentially "crowd-sourced" predictions then represent a powerful ensemble of model, a so-called collective AI, that is not longer confined to the limitations of single-entity control or perspective. Collective AI here refers to the synergy of multiple AI models working together, each bringing its unique strengths to create a more comprehensive and accurate solution. This approach leverages the diversity of data, methodologies, and perspectives, leading to enhanced problem-solving capabilities. Such collective AI is particularly potent in tackling complex, multifaceted problems where single-model approaches fall short.

1.3 On the Role of Blockchain

1.3.1 Blockchain as an Enabler for Decentralized AI

Blockchain technology emerges as a pivotal enabler for decentralized AI, primarily due to its ability to ensure fair rewards (i.e. enforced by smart contracts) for diverse model contributions in a trustless environment. A decentralized ledger can eliminate the need

for central coordination, thereby fostering a more equitable and transparent ecosystem. In such a setup, contributors of different AI models can retain ownership of their training data, model architecture and training methods, while being fairly rewarded based on the value they provide.

1.4 Our Core Values & Objectives

At the heart of *Quatern.ai* lie the foundational principles of collaboration, innovation, and transparency. We are dedicated to the democratization of AI, with the goal of creating an ecosystem where both individuals and organizations can contribute to, and reap the benefits of, a collective intelligence. Our mission is to forge a platform that not only stands at the forefront of technological advancements but also upholds ethical standards and prioritizes the needs and contributions of its users.

We stand by the conviction that a smart-contract-governed AI platform, rooted in the ethos of decentralization, offers a more equitable, inclusive, and potent platform for AI advancements than the prevalent centralized models. Our vision is to cultivate a future where the most pressing data science challenges are addressed collectively. In this envisioned future, every participant has the opportunity to contribute, receive recognition and fair compensation for their contributions, and maintain ownership of their intellectual property. Furthermore, this ensures universal access to state-of-the-art models developed within this decentralized framework.

The potential applications of this revolutionary approach span across all sectors and industries. We eagerly anticipate the diverse and impactful problems our community will tackle. As we stand at the precipice of this exciting journey, our enthusiasm for the transformative possibilities of decentralized AI remains boundless. The adventure is only beginning, and we invite you to join us in shaping the future of artificial intelligence.

2 System & Protocol Design

2.1 Overview

In order to give a first overview of our system and protocol design, we define the following key participating roles and their respective purposes/tasks:

- **Challenge providers** propose the problems to be solved by providing labeled data (through which the learning objective is defined). The data may come in different forms as discussed in Section 2.3.

- **Model developers** provide predictions for a the problem posed by a specific challenge. More concretely, this consists of

  1. Designing and implementing a model architecture. In order to help developers get started and bootstrap the process, *Quatern.ai* will be providing open-sourced baseline implementations for different problem types that can also be build on by non-programmers.

  2. Training the resulting model on the provided dataset.

  3. Deploying the trained model by either a) using *Quatern.ai*'s one-click deployment service, or b) self-hosting your model (instructions and tooling for this will be provided). While a) is a simple and straightforward deployment solution, b) gives advanced developers full control and in particular allows them not having to trust any third party with keeping their model architecture private.

- **Users** aim to use the finished (collaborative) model to generate predictions for their own problem inputs.

As becomes apparent, the key step lies in turning a set of predictions inferred by the

individual models provided by developers into one powerful combined prediction for each data point. This needs to be done in a way that

1. produces as good of an ensemble model as possible,

2. rewards fairly according to their contribution to this ensemble, and

3. does not provide any incentive for predictions to be maliciously manipulated by either model developers or challenge providers.

The next section outlines how we can simultaneously achieve all three of these objectives.

2.2 "Out of many, one."

To describe the model collaboration process, we look at the following four phases that each challenge/problem undergoes on *Quatern.ai*:

1. **Challenge proposal phase:** This phase is initiated through a challenge provider that proposes a new problem by providing a labeled dataset and a reward budget given in *$QAI* tokens. The dataset must be split into `train`, `test1`, `test2`, and `test3`. Only `train` is made public initially, for `test1` and `test2` only the unlabeled input data is published, and `test3` remains fully private (for reasons to become clear later). The reward budget is locked into a smart contract.

2. **Competition phase:** During the competition phase, developers design models and train them on the public `train` dataset. Submissions are evaluated based on predictions on `test1` with respective scores being showcased in a live public leaderboard.

3. **Evaluation phase:** Next, after the competition time window has ended, the developers' models are evaluated on `train2` (to avoid overfitting during competition phase). The resulting performance score determines both the individual model's weight in the final ensemble model, as well as the fraction of the reward budget paid out to its developer from the *$QAI* amount initially locked in the smart contract.

4. **Usage phase:** Finally, the completion of the evaluation phase launches the open-ended usage phase, in which users can access the resulting ensemble model for

computing predictions on their own data. To do so, users provide their input data and send a fee paid in QAI to the smart contract. This initiates the individual models to make their predictions which are then gathered and weighted according to the weights determined in the evaluation phase. The user then receives their prediction. The number of predictions a model has participated in with a certain weight is maintained in the smart contract from which developers can then unlock their respective share of QAI fee. In order to prevent developers from malicious behavior during usage phase (e.g. replacing their good model with a cheap model), model weights are continuously re-evaluated using data point from `test3` that are randomly provided instead of user inputs. The reason `test3` needs to initially remain fully private is hence to avoid that developers know when their models are being re-tested during usage phase.

2.3 Applications & Real-world Use-cases

Potential applications of ML/AI for data analysis are ubiquitous these days and can essentially be found across nearly all disciplines and sectors. To provide some examples, we briefly discuss real-world use-cases across the following three example areas.

2.3.1 Healthcare & Medicine

Predictive machine learning models may soon (and in some cases have already) outperform human doctors and other human medical experts on various tasks ranging from patient diagnosis and estimation of treatment outcomes to personalized medicine based on genetic data, and AI-driven drug discovery processes. One aspect that makes applications running on medical data particularly interesting in the context of decentralized collaborative learning is that such data is oftentimes subject to strict privacy constraints. In this case, individual models could be trained locally without any patient data leaving the hospital, and predictions could then be fed into an ensemble such that future predictions may still draw from learning structures globally. Indeed, such use-cases have already been widely studied and discussed in the context of traditional (i.e. non-decentralized) federated learning, see e.g. [Xu et al., 2021].

2.3.2 Financial Markets

The integration of ML/AI into financial markets has revolutionized the filed: From algorithmic trading to credit risk assessment and fraud detection, the potential applications are vast and varied. A notable pioneer in the realm of decentralized AI within finance is Numerai [Craib et al., 2017], which operates as a decentralized hedge fund by crowdsourcing predictive financial models. Participants submit predictions that are aggregated into a meta-model, which then informs investment decisions. While Numerai focuses on the specific application of creating a decentralized hedge fund, *Quatern.ai* aims to establish a more general framework for decentralized AI applications. Unlike Numerai's model, which is particularly suited for time-series data where the future labels (e.g., asset prices) are unknown, *Quatern.ai* method is designed to be applicable in a broader scope and across a wider range of data types, including non-time series datasets.

Moreover, *Quatern.ai* framework emphasizes security and privacy, allowing participants to contribute without exposing sensitive financial data. This is particularly relevant in the financial sector, where data privacy and protection are paramount. This opens up new avenues for innovation in financial analytics, where the collective insights derived from multiple models can lead to more robust, accurate, and comprehensive financial predictions and analyses.

2.3.3 Agriculture & Environmental Modeling

The agricultural sector stands on the brink of a technological revolution, with decentralized AI poised to play a vital role [Žalik and Žalik, 2023]. By harnessing decentralized data sources, AI models can optimize crop yields, minimize waste, and manage resources more sustainably. For instance, precision agriculture models utilize data from various sensors across farms to make informed decisions about planting, watering, and harvesting, tailored to the specific needs of each plot. This not only boosts productivity but also reduces the environmental impact of farming practices. Moreover, disease and pest prediction models represent a critical application of AI in agriculture. These models can proactively identify the risk of outbreaks, enabling farmers to take preventative measures or use targeted treatments rather than resorting to widespread pesticide use. Despite saving costs such measures may also mitigate the environmental damage caused by excessive

pesticide application.

In the realm of environmental modeling, utilizing data from a decentralized network of sensors and satellites, AI models can forecast weather events, track climate change impacts, and model ecosystem dynamics with unprecedented accuracy. The decentralized nature of these data sources and the collaborative approach to model building means that predictions can draw from a rich, diverse dataset as well as model architecture composition, leading to more robust and comprehensive models. This collective intelligence approach not only enhances the predictive power of environmental models but also democratizes access to these insights, allowing communities worldwide to prepare for and respond to environmental challenges more effectively.

3   *$QAI* Token

Note: The *$QAI* **tokenomics** and **distribution mechanisms** will be disclosed at a later stage during Phase 1 of the project (see our roadmap in Chapter 4 below) and added here accordingly.

3.1   Utility

The *$QAI* token is at the heart of the *Quatern.ai* ecosystem, designed to incentivize and facilitate the decentralized AI platform we envision. Its utility spans across rewarding both model performance and deployment, as well as empowering the community with governance rights. Here's how the *$QAI* token is instrumental in achieving these goals.

3.1.1   Rewarding Contributions

In the *Quatern.ai* ecosystem, the *$QAI* token acts as a direct incentive for model developers. During the training and competition phases, developers submit their AI models to tackle the respective challenge. As described in Chapter 2, the performance of these models is determined in the evaluation phase, and developers are rewarded according to their model's relative conribution to the resulting ensemble. This system sets the incentive for achieving the best possible ensemble model while fairly rewarding developers for their contribution.

3.1.2   Facilitating Model Usage

During the "Usage phase", a deployed ensemble model can be invoked by users in exchange for *$QAI* tokens. This fee is distributed to the underlying individual model deployers

compensating them for any operational and compute cost of hosting such a model. The use of *$QAI* tokens for such compensation ensures a seamless, integrated experience within the ecosystem, fostering a self-sustaining economic model.

3.1.3   Governance and Steering the Project's Direction

Beyond usage for rewards and fees, *$QAI* tokens grant holders governance rights within the *Quatern.ai* ecosystem. Once the platform's development moves past the initial stages, token holders will have the opportunity to propose and vote on key decisions regarding the project's further development, such as protocol upgrades, feature additions, and the allocation of resources in a community treasury. This participatory approach ensures that *Quatern.ai* evolves in a direction that reflects the values and priorities of its community, making *$QAI* holders integral to the decision-making process.

The multifaceted utility of the *$QAI* token underscores its central role in empowering the vision for a decentralized, collaborative AI platform. By aligning incentives across all stakeholders, the *$QAI* token facilitates a vibrant ecosystem where innovation, participation, and governance converge to shape the future of decentralized AI.

4 The Road Ahead...

4.1 Phase 1: Foundation (Q1-Q2 of 2024)

The initial phase of our journey, spanning the first half of 2024, is focused on establishing a solid foundation and laying the groundwork for our project.

- **Whitepaper Publication:** We will introduce *Quatern.ai* with the publication of the first version of this whitepaper, detailing the vision, framework, and anticipated impact of our project. This document will be continuously extended and improved, and will serve as a guiding manifesto for our venture.

- **Tokenomics Disclosure:** The strategy for token distribution and the utility of our tokens will be transparently shared with our community. This step is crucial for ensuring fair launch and long-term alignment of interests among all *Quatern.ai* stakeholders.

- **Community Cultivation:** A robust community is vital for the success of *Quatern.ai*. We will engage and grow our community, fostering active participation and support.

- **Protocol Development Journey:** Development will be underway to ensure the creation of robust smart contracts, a stable backend infrastructure, and a user-friendly interface.

4.2 Transition & Token Genesis (Mid 2024)

As we progress into mid 2024, *Quatern.ai* enters a crucial phase marked by the genesis of our token.

- **Token Inception Event:** The launch of our token will be initiated through a public sale, offering early community members the chance to be pioneers in the *Quatern.ai* ecosystem.

- **Testnet Showcase:** A demonstration of our protocol's capabilities and potential will be presented in a public beta on our Testnet.

- **Marketing & Engagement:** Our efforts to market *Quatern.ai* will intensify, aiming to expand our reach and deepen community engagement, paving the way for widespread adoption.

4.3 Phase 2: Platform Evolution (Q3-Q4 of 2024)

In the second half of 2024, *Quatern.ai* will focus on evolving and finalizing the platform.

- **Platform Development:** The protocol, infrastructure, and user interface will be finalized to ensure a robust and effective platform.

- **UX Refinement:** Through comprehensive user testing, the user experience will be continually refined to ensure intuitiveness and ease of use.

- **Mainnet Deployment:** The official launch of *Quatern.ai* on the mainnet will mark a significant milestone, bringing our decentralized AI platform to fruition.

- **Feature Extensions:** Additional features and components will be announced along the way and introduced post-launch to enhance and complement the core functionalities and keep evolving the platform.

- **Strategic Marketing:** Targeted marketing campaigns will be executed to solidify *Quatern.ai* market position and encourage widespread adoption.

Bibliography

[Breiman, 1996] Breiman, L. (1996). Bagging predictors. *Machine learning*, 24:123–140.

[Christian, 2020] Christian, B. (2020). *The alignment problem: Machine learning and human values*. WW Norton & Company.

[Craib et al., 2017] Craib, R., Bradway, G., Dunn, X., and Krug, J. (2017). Numeraire: A cryptographic token for coordinating machine intelligence and preventing overfitting. *Retrieved*, 23:2018.

[Xu et al., 2021] Xu, J., Glicksberg, B. S., Su, C., Walker, P., Bian, J., and Wang, F. (2021). Federated learning for healthcare informatics. *Journal of Healthcare Informatics Research*, 5:1–19.

[Žalik and Žalik, 2023] Žalik, K. R. and Žalik, M. (2023). A review of federated learning in agriculture. *Sensors*, 23(23):9566.