

Bazaar AI

AI Data Marketplace on Base

Introduction

High-quality data has become one of the most valuable resources for training AI models. However, obtaining such data is often a significant challenge.

The acquisition of valuable data is frequently limited to large corporations and private deals, leaving individual data providers uncompensated and their data utilized without their consent.

Bazaar AI presents a decentralized solution that addresses these challenges. By compensating data providers and enabling the investment and ownership of data sets by anyone, the marketplace facilitates the sale and purchase of a broad range of verifiable and high-quality data. This approach democratizes access to valuable data, fosters innovation, and ensures that data providers are fairly compensated for their contributions.

Bazaar AI

Bazaar AI is a Decentralized AI Data Marketplace on Base for requesting, funding, and purchasing high-quality datasets.

The platform operates as follows:

Crowdfunding: Anyone can request a specific dataset and pool together funds to incentivize nodes to provide the requested data. Funds can be contributed in various cryptocurrencies such as ETH, USDC, etc.

Data Provision: Nodes compete to provide the requested data, which is then stored on IPFS and made queryable via The Graph. As data streams in, funds are distributed to the nodes providing the data.

Data Access and Ownership: Once the data is available, it can be accessed by the owners (those who contributed funds) and purchased by others on the marketplace. The owners of the data receive profits from the sales.

Ensuring Quality of Data

To ensure the quality of data, nodes are required to stake tokens. If the data provided is deemed inaccurate, the node's tokens can be slashed. Zero-knowledge proofs can be used to verify the accuracy of the data without exposing private information. However, if the data is found to be inaccurate, it can be exposed by the owners without any loss.

Staking and Bidding

Nodes act as middleware that brings data on-chain and determines the revenue share to stakers. This helps nodes win contracts from the marketplace. The bidding process is typically decided by a dual system of reputation and tokens staked. There is no price bidding; nodes receive the pooled funds. If no high-quality nodes apply, the pooled funds can be adjusted through free-market mechanics.

Integrating and Compensating Data Providers

Node operators provide a simple API to integrate data sources. They are responsible for searching, finding, and integrating high-quality data providers, as well as designing the compensation framework. A smart contracting system is available to automatically distribute funds to data providers if they choose to use it as a base. Data is relayed from the source to the node via the API, where it is chunked, encrypted, and access control enabled via DID before being stored in

IPFS. Only data providers, node operators, data owners, and purchasers via the marketplace can decrypt and view the data.

AI Data Marketplace Platform

The AI Data Marketplace will be launched on Base, a superchain ecosystem built on top of Ethereum. By leveraging Base, our platform will benefit from the speed, stability, and robust infrastructure provided by the superchain, while maintaining the security characteristics of Ethereum.

To expedite development, we are utilizing code from an abandoned DeFi project, MyBit. After extensive due diligence, our technical team determined that the code from MyBit, which was an IoT crowdfunding and investment marketplace that completed extensive code audits, can be easily modified for an AI data marketplace. This allows us to significantly reduce development time by over a year and focus solely on the node software. As a result, we can quickly go to market with minimal resources and external investment.

By utilizing existing, audited code and launching on a secure and scalable platform, we aim to provide a reliable and efficient solution for data acquisition and monetization that can be brought to market quickly.

Use Cases

The AI Data Marketplace is designed to facilitate the acquisition, monetization, and sharing of high-quality datasets for various AI-specific use cases across different industries. Here are some examples of how the platform can be utilized:

- **AI Model Training:** AI and machine learning models require large amounts of high-quality data for training. The AI Data Marketplace enables data scientists and AI developers to access and purchase

relevant datasets for training their models, improving their accuracy and performance.

- Natural Language Processing (NLP): Developers working on NLP applications such as chatbots, virtual assistants, and language translation tools can access diverse datasets containing text, speech, and language-related data to improve their models' understanding and generation of human language.
- Computer Vision: Researchers and developers working on computer vision applications, such as object detection, facial recognition, and autonomous vehicles, can access large datasets of images and videos to train and improve their models' ability to interpret visual data.
- Predictive Analytics: Companies can use the AI Data Marketplace to access diverse datasets related to market trends, customer feedback, and other relevant factors to develop predictive models that can anticipate future outcomes and inform decision-making.
- Sentiment Analysis: Social media platforms, marketing agencies, and other organizations can access datasets containing anonymized social media posts, reviews, and other user-generated content to develop AI models that can analyze and interpret public sentiment towards products, brands, and topics.

Token [BAZ]

The AI Data Marketplace will have a total new supply of 400 million tokens, distributed as follows:

* 45% airdropped to MYB holders

- * 35% allocated to the community airdrop
- * 11% allocated to the team, with a 6-month cliff and 1-month linear vesting for 2 years
- * 9% allocated to the DAO treasury for future fundraising

Inflation will be set at 8% and will halve every year. This inflation will be used to subsidize node revenue as transactions continue to grow.

The smart contracts making up the AI Data Marketplace will be governed by token holders via a simple DAO structure. Token holders will have the power to allocate funding, adjust fees, inflation, and other parameters of the contracts. This ensures that the platform is decentralized and community-driven.

Airdrop Roadmap

The AI Data Marketplace will follow a phased approach for its launch and airdrop distribution.

Phase 1 Airdrop: Snapshot of MYB holdings on Ethereum

- * A snapshot of MYB holdings on Ethereum will be taken.
- * Airdrop Phase 1 will be distributed to MYB holders to reward them for using the code from their project as the foundation for the data investment, marketplace, and revenue distribution engine.
- * These tokens will become immediately claimable at the start of phase 2: community airdrop.

Phase 2: Community Airdrop

- * Phase 2: community airdrop will launch with multiple challenges throughout the test-net period.
- * Phase 2 will begin immediately at the completion of Phase 1 airdrop when tokens are released for claiming.

- * Airdrop points will be earned by staking on nodes to signal demand for different datasets.
- * Staking ETH in the data node contract will earn 1 point daily per dollar value of ETH.
- * Staking BAZ in the data node contract will earn 3.5 points daily per dollar of value of BAZ.
- * Running a data node will earn 5% of the total points awarded to depositors in your data node contract per day.
- * Referring friends will earn 15% of their points daily.
- * Airdrop allocation will be calculated by dividing your points by total points and multiplying by 140 million BAZ.
- * All airdrop tokens will be released without vesting at launch.

By following this roadmap, we aim to incentivize community participation, signal demand for different datasets, and ensure a fair and wide distribution of tokens.