# Chainbase: The Decentralized Omnichain Data Network for Open AGI Economy

Chainbase Labs

November 28, 2024

**Contents**

# 1 Chainbase Overview

Artificial Intelligence (AI) is transforming industries by automating tasks, enhancing efficiency, and enabling data-driven decision-making. AI's potential is driven by its capacity to evaluate enormous datasets and extract insights, enabling well-informed judgments in a variety of industries, including healthcare, finance, and transportation. Through customization and user-friendly technological interfaces, the incorporation of AI into everyday activities further enhances user experiences.

However, the exponential demand for high-quality, structured, and trustworthy data to fuel AI systems presents significant challenges. Addressing these challenges requires innovative solutions that prioritize data accessibility, integrity, and usability. Chainbase tackles these issues with its decentralized data network, designed to break down barriers and ensure that data meets the needs of modern AI systems. By leveraging blockchain technology, Chainbase ensures secure, transparent, and efficient data management, empowering developers to create useful AI solutions while incentivizing contributions in a knowledge-based economy.

# 2 Priorities for Chainbase

The Chainbase network is founded on the principle of decentralization, ensuring that data is neither controlled nor processed by a single entity. This decentralized approach aligns with Chainbase's four core priorities: Open Source, Collaboration, Incentivization and AI Readiness.

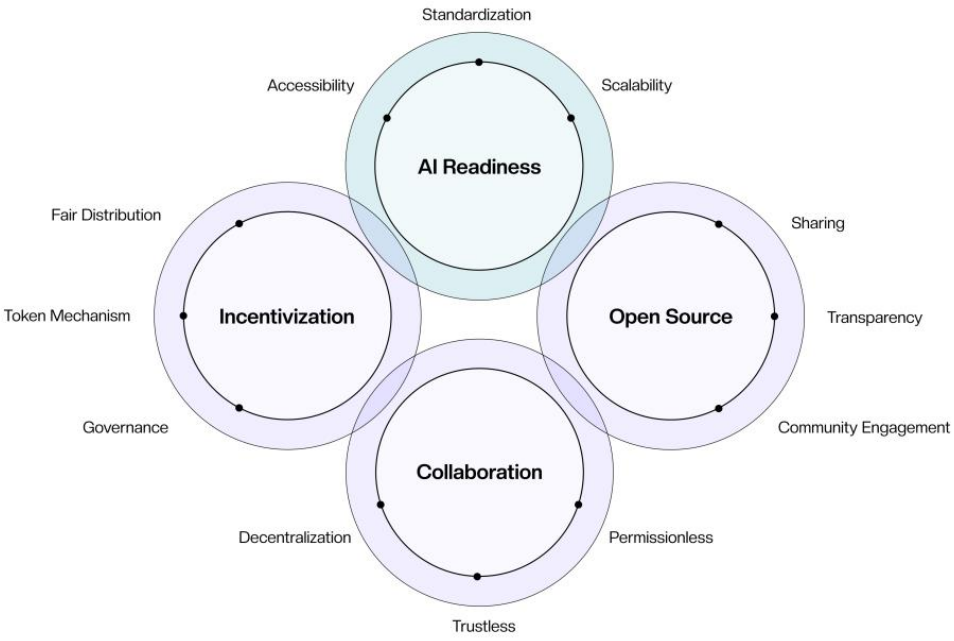


Figure 1: Four pillars of Chainbase Network

## 2.1 Open Source

In an era driven by data, transparency and openness are essential for fostering innovation and trust. Open-source software provides developers with the tools to iterate, improve, and create in groundbreaking ways. As the AI revolution unfolds, projects and protocols that openly share their frameworks with a skilled technical community will lead the charge in shaping the future. Chainbase is committed to open-sourcing key developments and cultivating an engaged community of developers. By empowering contributors to build and expand on its innovations, Chainbase ensures sustained growth, ingenuity, and accessibility for blockchain technologies well into the future.

## 2.2 Incentivization

A thriving, community-driven ecosystem requires a fair and transparent incentivization model that rewards all contributors and promotes innovation. Chainbase leverages blockchain technology to meticulously record data operations and contributions, providing clear attribution to data providers, AI model developers, and other network participants. This transparency ensures equitable compensation and supports the ethical development of AI. Chainbase's architecture actively incentivizes contributions across the network, whether through securing the decentralized infrastructure, providing valuable datasets, validating data, or sharing computational power. The $C token is central to this system, enabling secure and transparent economic circulation while encouraging contributors to share resources and expertise for the collective benefit.

## 2.3 Collaboration

Collaboration lies at the heart of Chainbase's vision for a data-driven ecosystem. By eliminating reliance on centralized authorities, the network drives an inclusive and cooperative environment where participants can share and benefit from collective knowledge. Chainbase recognizes that monopolistic approaches cannot meet the demands of the advancing AI era. Through its permissionless framework, the network maximizes collaborative potential across roles, empowering participants to define and share data-driven insights tailored to emerging challenges and opportunities.

## 2.4 AI Readiness

High-quality, standardized data is the lifeblood of AI development. Chainbase is dedicated to establishing a unified standard for diverse data sources, ensuring hygiene and compatibility with advanced AI applications. At the core of this initiative are Manuscripts, which streamline data formatting and standardization, making it seamlessly accessible for AI systems. Designed to meet the demands of data-intensive AI projects, Chainbase's

network serves as a foundational infrastructure for developers, enabling them to harness vast, high-quality datasets to drive innovation. By providing reliable data at scale, Chainbase enhances the efficiency and intelligence of AI models, enabling them to achieve greater accuracy and functionality per parameter. This critical improvement empowers the development of next-generation AI technologies, driving innovation and expanding their potential applications.

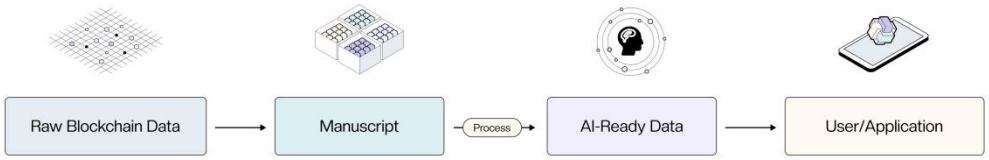


Figure 2: AI Readiness Data

# 3 Chainbase Four-Layer Architecture

The Chainbase network operates on a sophisticated four-layer architecture designed to provide secure, efficient, and scalable data management and processing capabilities. Each layer serves a distinct purpose, seamlessly integrating to ensure a decentralized environment optimized for collaborative knowledge sharing, robust execution, and consensus-driven validation and high-quality data accessibility.

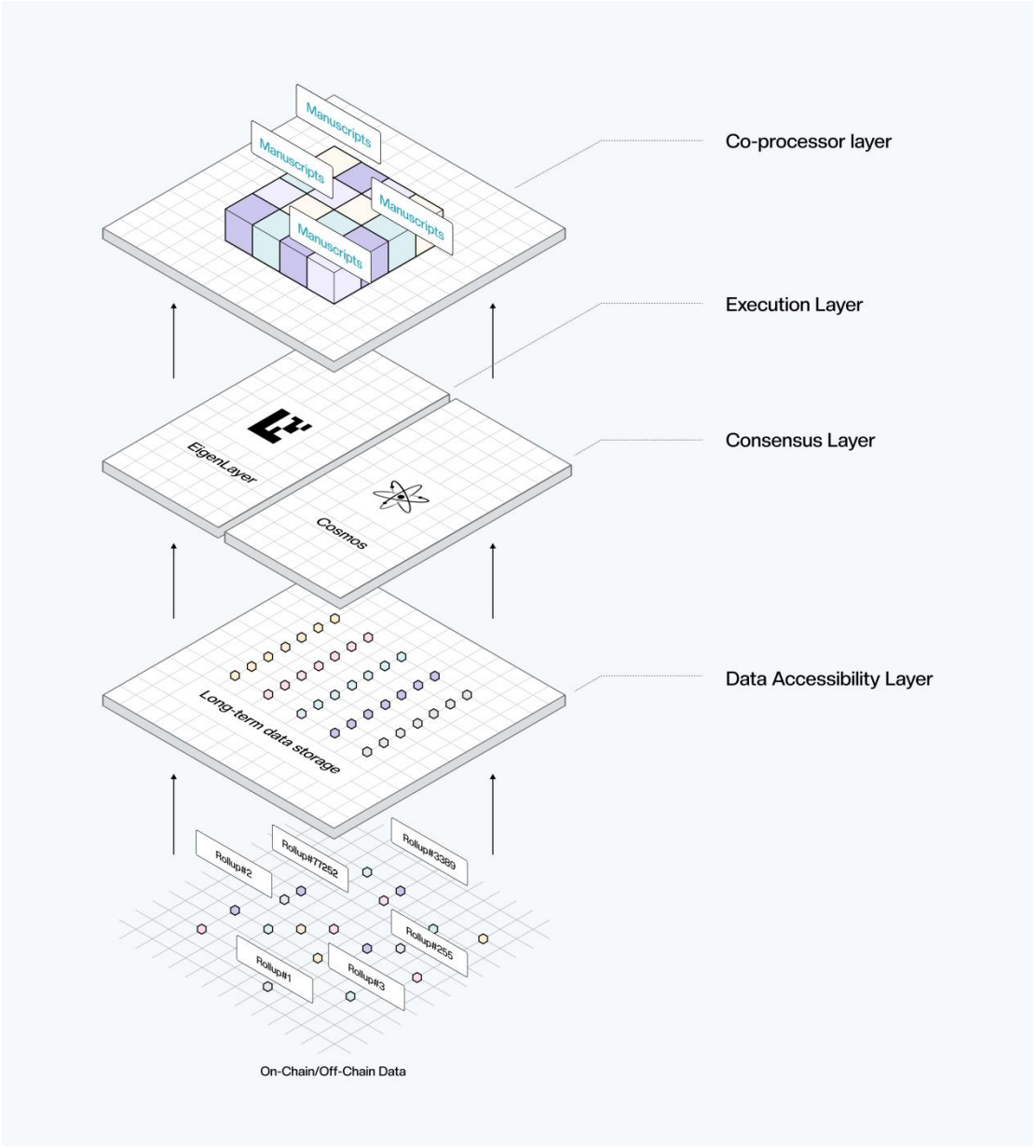


Figure 3: Four Layer Structure of Chainbase Network

## 3.1 Co-Processor Layer

The Co-Processor Layer is designed to cultivate collaboration by enabling users to contribute data processing expertise and AI knowledge to the network. Central to this layer is the concept of "Manuscripts," which serve as programmable scripts defining data transformation workflows.

Manuscript is a crucial element of the Chainbase ecosystem, acting as the language for defining and executing data processing tasks. They allow developers to standardize data formats and processing tasks, making raw data AI-ready. This layer also enables

interoperability by seamlessly integrating data from various blockchains and off-chain sources. Additionally, contributors can monetize their knowledge by transforming their work into valuable assets, with Manuscripts treated as tradable resources within the network. The $C token plays a critical role here, acting as the currency for payments, settlements, staking, and governance. The Co-Processor Layer not only facilitates collaborative knowledge-sharing but also establishes a transparent economy that rewards participants for their valuable contributions. By enabling the assetization of expertise and promoting fair compensation, this layer drives innovation and enhances the overall utility of the network.

## 3.2 Execution Layer

The Execution Layer acts as the computational powerhouse of Chainbase, responsible for executing Manuscripts and managing large-scale data processing tasks. At its core is the Chainbase Virtual Machine (CVM), a custom-built virtual environment optimized for processing Manuscripts and executing data workflows. This layer is designed to handle vast volumes of data efficiently by employing advanced processing techniques such as data parallelism and task parallelism. Data parallelism enables simultaneous processing of different chunks of datasets, while task parallelism allows multiple tasks to be executed concurrently, optimizing resource utilization. The Execution Layer also integrates distributed verification protocols, such as Eigenlayer, to enhance decentralization and security. Node operators play a vital role by providing the computational resources required for smooth operations, with their contributions rewarded based on workload and performance. By ensuring scalable, efficient, and secure processing of Manuscripts, the Execution Layer empowers developers to execute complex AI tasks while maintaining high performance and reliability.

## 3.3 Consensus Layer

The Consensus Layer underpins the security and integrity of the entire network. It ensures that all transactions and data states are validated and agreed upon by the network's participants. Chainbase utilizes the CometBFT consensus algorithm, a Byzantine Fault Tolerant (BFT) mechanism that guarantees resilience against network failures and malicious attacks. This algorithm provides instant finality, ensuring near-instant data freshness and seamless updates across the network. The Consensus Layer employs a Delegated Proof of Stake (DPoS) system, where validators, selected based on $C token stakes, maintain blockchain integrity, validate data operations, and ensure consistency. Delegators further enhance network security by staking their $C tokens with trusted validators, reinforcing the system's economic resilience. Validators are rewarded for their critical role in ensuring the accuracy and stability of the network. By aligning economic incentives with network security, the Consensus Layer maintains a robust and trustworthy decentralized framework.

## 3.4 Data Accessibility Layer

The Data Accessibility Layer forms the data foundation of the Chainbase network. It is responsible for collecting, verifying, and storing both on-chain and off-chain data which respond to the data processing functionalities of Manuscripts. On-chain data, stored directly within blockchains, is immutable, transparent, and secure, making it ideal for recording transaction histories, staking information, and metadata. Off-chain data, on the other hand, resides in decentralized storage systems, addressing scalability and privacy challenges. Examples of off-chain data include large raw datasets, programming code, and complex AI models. This layer ensures data integrity by sourcing information from a decentralized network of providers, preventing control or manipulation by any single entity. By utilizing advanced cryptographic techniques like Zero-Knowledge Proofs (ZKP), it validates data sources without exposing sensitive details, while consensus mechanisms ensure the trustworthiness of the data before it is permanently stored in decentralized systems. This approach eliminates data fragmentation, provides real-time access to verified information, and ensures high-quality data readiness for AI applications.

## 3.5 Integration and Functionality

The seamless interaction between these four layers aims to ensure Chainbase's ability to deliver high-quality, AI-ready data in a decentralized and secure manner. The Data Accessibility Layer provides verified and unified datasets, while the Co-Processor Layer transforms this data into standardized formats through collaborative efforts. The Execution Layer efficiently processes and executes workflows, and the Consensus Layer secures the entire network through resilient and transparent validation. Together, these layers empower all network participants to collaborate effectively, driving innovation in AI and data-driven applications within a decentralized ecosystem.

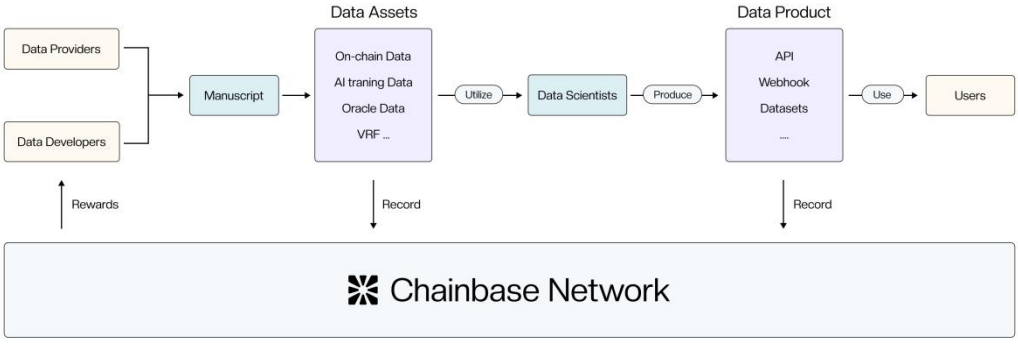


Figure 4: User journey of Chainbase Network

# 4 Chainbase Network Participants

The Chainbase network is built on the contributions of a diverse ecosystem of participants, each playing an essential role in ensuring its growth, stability, and functionality. These participants—developers, node operators, validators, and delegators—collaborate within the network's architecture to maintain a seamless, decentralized, and secure environment.

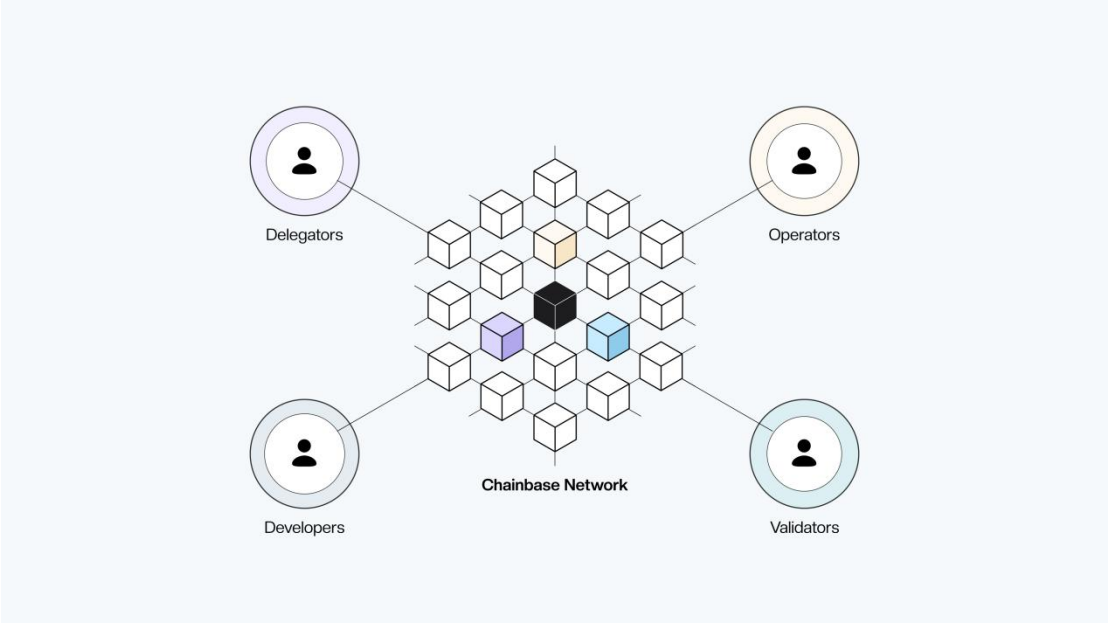


Figure 5: Participants of Chainbase Network

## 4.1 Developers

Developers are at the core of Chainbase's innovation, engaging with all layers of the network to create, enhance, and optimize data-driven applications. At the Data Accessibility Layer, developers access raw and processed datasets to power their Web3 applications and AI models. The Co-Processor Layer enables developers to contribute their expertise by creating and sharing "Manuscripts," which are programmable scripts defining data processing workflows and transformations. These Manuscripts help standardize data formats, making them suitable for advanced AI applications while allowing developers to monetize their contributions. Lastly, developers utilize the Execution Layer to run these Manuscripts within the Chainbase Virtual Machine (CVM), a specialized environment that efficiently processes large-scale data. By interacting with the network's layers, developers drive Chainbase's ecosystem forward, fostering innovation and expanding its utility.

## 4.2 Operators

Node operators are responsible for providing the computational power necessary to sustain the network's Execution Layer. They play a critical role in processing Manuscripts and

executing data workflows, ensuring the smooth operation of AI and data processing tasks. Node operators manage the heavy lifting of computation, leveraging their resources to maintain network performance and efficiency. Their contributions are directly linked to workload and performance, with rewards distributed in proportion to their efforts. By dedicating their infrastructure to the network, node operators ensure scalability and reliability, enabling Chainbase to handle complex and large-scale data operations.

## 4.3 Validators

Validators are the guardians of network integrity, maintaining security and ensuring consensus through their work in the Consensus Layer. Their primary role is to validate data operations, verify transactions, and uphold the consistency of the blockchain's state. Validators rely on the CometBFT consensus algorithm, which ensures Byzantine Fault Tolerance and provides instant finality for data and transactions. By participating in the consensus mechanism, validators help secure the network against failures and malicious attacks. Validators receive block rewards for their contributions, incentivizing them to maintain high standards of accuracy and reliability. Their critical role in safeguarding the network's decentralized framework ensures Chainbase's long-term stability.

## 4.4 Delegators

Delegators play a supporting role in the Chainbase ecosystem by enhancing the network's economic security. While they do not directly participate in tasks like computation or validation, delegators contribute by staking $C tokens with trusted validators and node operators. This staking mechanism bolsters the resilience of both the Execution Layer and the Consensus Layer, incentivizing active participation and strengthening the network's economic incentives. In return, delegators earn a share of the rewards generated by validators and node operators. Their contributions enable a more decentralized and secure network, ensuring sustained participation and collaboration across the Chainbase ecosystem.

# 5 Introduction to $C

The $C token serves as the backbone of the Chainbase ecosystem, facilitating interactions, incentivizing participation, and ensuring the network's operational efficiency and sustainability. With the initial economic circulation of 1,000,000,000 tokens, $C's utilities are designed to align the interests of various participants within the Chainbase network, nurturing a collaborative and robust decentralized data economy.

## 5.1 Network Fees

$C is used to pay fees for accessing and querying datasets processed through Manuscripts. These fees support the operational infrastructure by compensating network participants, including Operators and Validators. A significant portion of these fees is allocated to the original Manuscript creators, Operators and Validators, incentivizing them for contributing high-quality data query systems and data knowledge.

## 5.2 Incentive

Node Operators and Validators play crucial roles in maintaining Chainbase's infrastructure and are incentivized through dedicated $C token mechanisms. Node Operators, who provide the computational resources that power Chainbase, receive rewards from dedicated token pools distributed over time, fostering network scalability and encouraging the continuous development of processing capacity. Validators, responsible for maintaining network consensus and data integrity, are compensated with block rewards in $C. This comprehensive reward system aligns the interests of both Node Operators and Validators with the network's long-term stability, while encouraging active participation in securing and scaling the network infrastructure.

## 5.3 Staking and Delegation

Staking is a fundamental aspect of the $C ecosystem, requiring Validators and Operators to stake tokens as a prerequisite for participation and reward eligibility. Delegators, individuals who do not directly operate within the network, can strengthen its security and resilience by staking $C tokens with Validators and Operators. In return, Delegators share in the rewards earned by these participants, fostering a collaborative and economically secure environment.

## 5.4 Network Governance

$C Token delegators can participate in network governance and have the right to propose and vote on the direction of the network. They can either vote directly or entrust others to vote on their behalf, thereby enhancing the transparency, democracy, and independence of the network.

The $C token is the lifeblood of the Chainbase ecosystem. Through its carefully designed tokenomics, $C not only ensures a sustainable and secure network but also drives innovation and growth in the decentralized data and AI landscape. It goes beyond being a medium of exchange, serving as a foundational element that aligns participant incentives and supports the network's long-term vision.

# 6 Governance

Chainbase Network stands out for its thoughtfully constructed governance ecosystem, which merges the principle of decentralization with practical mechanisms for community participations. At its core, this governance model reflects Chainbase network's commitment to acknowledging community's contributions and ensuring sustainable growth as a whole.

## 6.1 Chainbase DAO

The Chainbase DAO represents the decentralization of the governance framework. This DAO structure allows community members to actively participate in decision-making, either directly or via elected representatives, ensuring inclusive representation of contributors. Unlike simple voting systems, Chainbase DAO is designed to support strategic resource allocation, balancing immediate developmental needs with long-term network goals. It also serves as a catalyst for innovation, incentivizing meaningful contributions from participants and ensuring the network remains dynamic and forward-thinking.

## 6.2 Chainbase Foundation

The Chainbase Foundation plays a crucial role in supporting the growth and stability of the network, especially during its early stages. Its main focus is on building and improving the network's infrastructure, connecting key stakeholders, and driving innovation. By providing guidance and resources, the Foundation ensures that the network has the tools it needs to succeed.

A major initiative led by the Foundation is the Manuscript development, an open-source effort aimed at strengthening Chainbase Network's core technologies. The Foundation prioritizes research in key areas like scalability, security, and decentralization, ensuring the infrastructure remains effective and ready to handle potential challenges.

## 6.3 A Dual-Track Governance Approach

By operating on this dual-track system—where the DAO drives community-driven decision-making and the Foundation anchors developmental priorities—Chainbase achieves a unique balance. The DAO ensures community autonomy, giving participants a voice and a stake in governance, while the Foundation provides clarity of purpose and efficient resource allocation during critical stages of growth.

Importantly, the Foundation is designed as a temporary entity, with its ultimate goal being to cultivate an ecosystem that can function entirely autonomously. This transition reflects Chainbase Network's commitment to true decentralization, ensuring the ecosystem evolves into a self-sustaining and community-driven entity

# 7 Looking forward

Chainbase is set to revolutionize the intersection of AI and blockchain, empowering participants to create, collaborate, and innovate within a decentralized, knowledge-driven ecosystem. The Manuscript framework enables standardized workflows to transform raw blockchain data into curated datasets, fostering precise and efficient AI training. These datasets, alongside a dynamic marketplace for data processing standards and algorithms, reduce barriers to AI adoption, allowing developers to monetize their expertise and access pre-built solutions for tasks like fraud detection or market prediction. By leveraging a decentralized network of Node Operators incentivized through the $C token, Chainbase democratizes AI training and execution, ensuring inclusivity and accessibility for smaller organizations and individual developers.

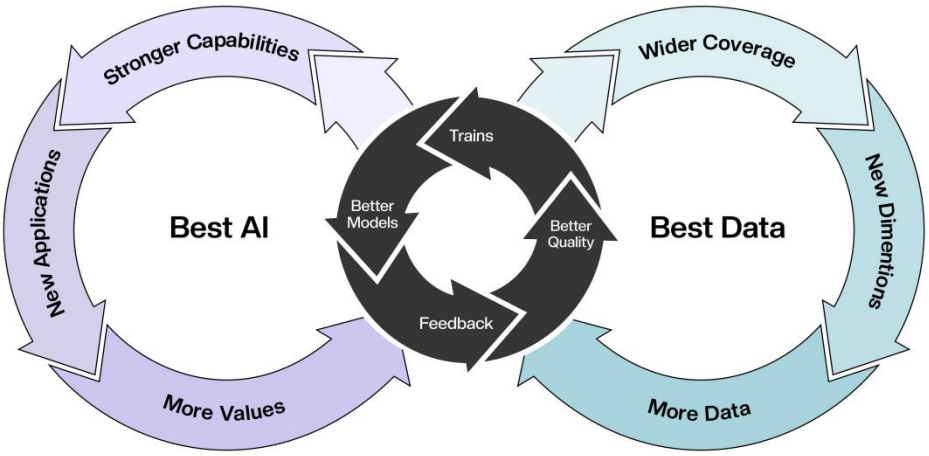


Figure 6: Flywheel: Best data, best AI

Blockchain integration within Chainbase ensures secure and transparent AI development, addressing critical concerns about data provenance, algorithmic integrity, and bias. The $C token underpins a knowledge-based economy that incentivizes data providers, developers, and participants to continuously contribute and innovate. This ecosystem bridges the gap between AI and Web3, enabling developers to enhance decentralized applications with advanced functionalities like intelligent automation and predictive analytics. By combining standardized data pipelines, decentralized computation, mature consensus mechanism and transparent AI practices, Chainbase is poised to become a transformative force in the convergence of AI and blockchain empowering the next wave of intelligent and decentralized solutions.

# Chainbase Glossary

- **Artificial Intelligence (AI):**

Technology that enables machines to mimic human intelligence, performing tasks like decision-making, problem-solving, and data analysis.

- **Blockchain:**

A decentralized and immutable ledger used to securely record and verify transactions across a distributed network.

- **Consensus Mechanism:**

A system used in blockchain networks to validate transactions and ensure all participants agree on the state of the ledger.

- **Decentralization:**

The process of distributing control and decision-making across a network, rather than relying on a central authority.

- **Validator:**

A participant in the blockchain network responsible for verifying transactions, maintaining integrity, and reaching consensus.

- **Node Operator:**

An individual or entity that provides computational resources to the Chainbase network for executing data workflows and running AI models.

- **Delegator:**

A participant who stakes their $C tokens with Validators or Node Operators to support the network and earn rewards.

- **$C Token:**

The utility token of the Chainbase network, used for staking, governance, payments, and incentivizing contributions across the ecosystem.

- **Staking:**

Locking up tokens as collateral to participate in the network's operations and earn rewards.

- **Manuscript:**

An executable script that runs on the Chainbase Virtual Machine (CVM). Manuscripts define standard schemas for converting raw data into structured datasets, ensuring compatibility and usability for AI applications.

- **Chainbase Virtual Machine (CVM):**

A custom-built execution environment that runs Manuscripts and performs sophisticated data processing and query operations efficiently.

- **EigenLayer:**

A decentralized protocol integrated into Chainbase to enhance network security and decentralization. It allows Validators and Node Operators to restake their assets across multiple protocols, ensuring robust execution and validation.

- **Data Network:**

The Chainbase ecosystem combining co-processing (data processing standards, $C tokens), execution (CVM, EigenLayer AVS), consensus (CometBFT, ABCI++), and data accessibility (roll-ups and network participants).

- **Knowledge Assetization:**

The process of turning contributions like Manuscripts and AI models into monetizable and tradeable assets.

- **Zero-Knowledge Proof (ZKP):**

A cryptographic method that proves the validity of data without revealing the data itself, ensuring privacy and integrity.

- **Data Accessibility Layer:**

The foundational layer in the Chainbase architecture responsible for collecting, verifying, and storing both on-chain and off-chain data.

- **Co-Processor Layer:**

The layer that enables collaboration by allowing users to contribute, monetize, and standardize data processing workflows.

- **Execution Layer:**

The computational layer responsible for executing Manuscripts and managing large-scale data processing tasks.

- **Consensus Layer:**

The security layer ensuring network consistency and agreement through mechanisms like Delegated Proof of Stake (DPoS) and CometBFT.

- **Smart Contract:**

A self-executing program with terms of agreement directly coded, used to automate processes on blockchain platforms.