DataR2E: Research and Prospects on the Value Release of Data Elements in Web 3.0

1st Shuning Sun
Hangzhou Qulian Technology Co., Ltd.,
Hangzhou Institute for Advanced Study,
University of Chinese Academy of Sciences,
Hangzhou, China
sunshuning23@mails.ucas.ac.cn

2nd Xi Ran
Hangzhou Qulian Technology Co., Ltd.,
School of Management Engineering and E-Business,
Zhejiang Gongshang University,
Hangzhou, China
ranxi169@163.com

3rd Shuchao Pang
School of Cyber Science and Engineering,
Nanjing University of Science and Technology,
Nanjing, China
pangshuchao@njust.edu.cn

*Xiaofeng Chen
State Key Laboratory of Blockchain and Data Security,
Zhejiang University,
Hangzhou Qulian Technology Co., Ltd.,
Hangzhou High-Tech Zone (Binjiang) Institute of Blockchain and Data Security
Hangzhou, China
chenxf.alfred@zju.edu.cn

*Yi Sun
Institute of Computing Technology,
Chinese Academy of Sciences,
Beijing, China
Hangzhou Institute for Advanced Study,
University of Chinese Academy of Sciences,
Hangzhou, China
sunyi@ict.ac.cn

2025 IEEE Global Blockchain Conference (GBC) | 979-8-3503-6642-6/25/$31.00 ©2025 IEEE | DOI: 10.1109/GBC60041.2025.11134480

***Abstract*—In Web 3.0, to achieve the continuous release of the value of data elements, we start with technologies related to data circulation and conduct research and analysis on the current data element market. We find that in the traditional model of data element circulation, the sharing and exchange of multi-source heterogeneous data are disconnected from the circulation and release of data element value. This results in poor data reusability, unclear data requirements, small market size, short lifecycle of data elements and data markets, making it difficult to motivate sellers to actively participate in building the data element market, and thus hard to achieve a positive cycle of data element value release. To address this, we have conducted a strategic analysis of the value release of data elements and designed a universal reference model, DataR2E, for the entire process of value release in the data elements market. Within the DataR2E model, we introduce the concept of data element bounty production, utilizing Web3 technology and data production tools to encourage sellers to actively provide the data elements needed by buyers. We envision using data production tools to build a bridge between buyers and sellers in the data elements market, promoting the sustainable development of the data element market, overcoming the mismatch in data elements expected by buyers and sellers, thereby resolving the issues of market presence without pricing and pricing without market presence, achieving sustainable development of the data elements market, and maximizing the potential of data element value release.**

***Index Terms*—blockchain, Web3.0, data element, value release, bounty production**

*Corresponding author

I. Introduction

The rapid development of data technology has smartened human social life, providing more autonomy and convenience, and has brought broad and profound impacts to multiple sectors such as economics, technology, healthcare, and social management. In the economic domain, the widespread application of big data analytics and artificial intelligence has revolutionized business operations and models. Through extensive analysis of business data, enterprises can better grasp market demands and predict business trends, thereby adjusting their economic strategies in a timely manner. As the demands for data applications evolve, the development of data technology is changing the way and scale of data input into production, with data gradually becoming a key element driving production in the respective technological and industrial environments [1]. Therefore, to highlight the importance of data in the digital economy, China in 2019 proposed for the first time to add data as a factor of production, using the term "data element" to refer to "data" in the context of productive forces and relations.

In the digital economy era, data has become an important resource for work and life. Numerous big data applications are built on the basis of data reuse, employing different analytical and processing operations on the same data to meet various needs of social production. For example, in the connected car industry, massive amounts of data generated by car sharing are collected by automotive service providers, who utilize data analytics to improve services and gain greater economic benefits [2]. Through various artificial intelligence technologies, the value of data is gradually being unlocked; however, this also raises issues related to privacy, security, and ownership rights [3].The leakage of personal information, invasion of privacy, and cyber threats have become the most significant challenges. Web3.0 has emerged in response; it represents the next generation of the internet, based on technologies like blockchain. It relies on smart contracts and similar technologies to restructure existing production relations, emphasizing a smarter, more decentralized, and less centralized internet experience. Compared to Web2.0, Web3.0 places greater emphasis on user control over data, highlighting openness, transparency, privacy, security, and enhanced interoperability [4].

On blockchains with smart contract capabilities [5], data elements can be integrated into automated business processes, reducing intermediaries, enhancing efficiency, and effectively lowering costs. In Web3.0, the decentralized philosophy of Web3 technology supports the creation and exchange of digital assets. The economic value of data elements can be directly converted into digital assets through technologies such as digitization, encryption, and smart contracts, facilitating more efficient value exchanges [6]. However, in economics, value exchange needs to form an effective closed loop to ensure the sustainable development of the market. For this purpose, this paper, set against the backdrop of Web3.0 and starting with technologies related to the release of data element value, analyzes the current development of the data element market. It introduces the concept of data element bounty production and devises a strategy for releasing the value of data elements. It envisions the use of points/tokens/coins as data bounty mediums in various industries to construct a universal reference model for the full process of data element market value release.

II. Related Technical Background

In Web3.0, the trusted circulation of data is one of the key means to promote the release of the value of data elements and is also the most important factor affecting the maximization of the value release of data elements. As a critical link in building a digital ecosystem, trusted circulation ensures the security, integrity, and availability of data in the digital economy. The "Trusted Data Circulation White Paper 2022" presents a universal four-layer architecture for trusted data circulation: the application layer, DaaS layer, data circulation control layer, and data circulation layer. It proposes a data circulation construction scheme that combines "distributed + centralized" approaches, using a centralized data custody model for general data and distributed data circulation services based on blockchain technology in Web3.0 for core data, and employing different data circulation schemes for different levels of data elements that are public, usable, and non-shareable [7].

To achieve the release of data element value, the construction of trusted data circulation solutions needs to include four core technological components: data integration, exchange and sharing, privacy computing, and blockchain. This section will start from the perspective of technological background and analyze the roles these four core technological components play in achieving secure sharing of data elements, effective utilization, and maximization of value release.

A. Data Integration

Data integration primarily involves data collection, data cleaning, and data storage. Data collection entails gathering data in various storage formats through different methods, ensuring the legality and compliance of data sources and strict control of data quality. Data cleaning is responsible for cleaning and correcting collected data that may have issues such as missing or duplicated entries, filtering out logical errors, and verifying data associations. Data storage requires the desensitization of data, setting access controls according to the needs of data usage, classifying and storing data based on the importance of the information, and encrypting data storage to meet security requirements.

B. Exchange and Sharing

Exchange and sharing are primarily responsible for the transmission and sharing of data elements, ensuring that data can flow securely and efficiently between different systems and organizations. Reliable data exchange and efficient data sharing are key issues in building data element markets. For example, in the healthcare sector, secure exchanges of health data can enable effective electronic transmission of clinical data between different health systems while ensuring its privacy and security.

In Web 3.0, with blockchain technology's decentralization, auditability, and transparency enabling the secure transmission of data elements. Combined with smart contract technology, data providers can quickly and efficiently control data sharing permissions, making data exchanges among participants more trustworthy and transparent without third parties. In the field of medical data sharing, most existing research, typically based on the cloud, cannot fundamentally address the issue of untrustworthy third parties and suffers from low efficiency in continuously generating and sharing data streams from sensors and other devices. To address these issues, scholars like Bingqing Shen

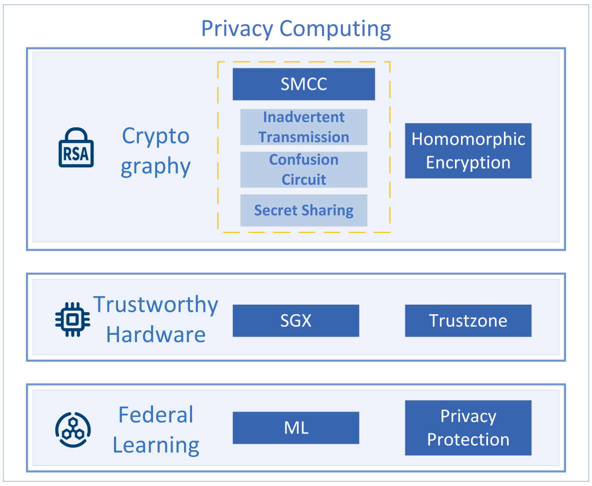

Fig. 1. Privacy computing.

have proposed utilizing the decentralized nature of blockchain, combined with P2P network technologies, to design an efficient data sharing blockchain, MedChain, which achieves higher efficiency while meeting the security requirements of data sharing [8].

The development of data exchange and sharing technologies has greatly ensured data security, thereby promoting the trustworthy circulation and value release of data elements, and providing a solid foundation for the development of the data elements market.

C. Privacy Computing

Privacy computing is the most effective means to address data privacy issues during the circulation of data elements. Relying on cryptographic principles and underlying hardware technologies, it allows data to maximize its value while ensuring privacy. Therefore, the development of privacy computing technology is particularly important for the trustworthy circulation and value release of data elements.

The development of privacy computing technology can be categorized into three main directions: cryptography, trusted hardware, and federated learning [9]. Technologies centered around cryptography mainly include secure multi-party computation and homomorphic encryption; trusted hardware is primarily led by Trusted Execution Environments (TEE), which rely on hardware to solve privacy protection issues; federated learning involves multiple participants combining data sources to collaboratively build models, aiming to enhance model performance and accuracy of output results through distributed machine learning.

D. Blockchain

Since the birth of Bitcoin in 2009 [10], blockchain technology, which underpins it, has rapidly evolved. In Web3.0, its applications have expanded into multiple domains including finance, supply chain, healthcare, and government services, among others [11]. This expansion has led to the emergence of various public, consortium, and private blockchain platforms, providing a solid decentralized foundation for the release of data elements' value.

Blockchain provides a decentralized, tamper-proof mechanism for recording data during the process of releasing the value of data elements. Its decentralized nature is particularly well-suited for establishing trust mechanisms in data circulation. Through its unique consensus algorithms and smart contract technologies, it facilitates the creation of transparent and secure data element trading platforms.

III. Research Related to Data Element Markets

The data element market is an organic system integral to achieving data integration, exchange sharing, and privacy protection. Building a robust data element market is essential for unleashing the value of data elements and promoting the high-quality integration and development of the real and digital economies.Currently, research related to the data element market mainly falls into three categories: economic models for data pricing, legal ownership studies for data confirmation, and supportive technologies for data transactions.

A. Data pricing

In 2016, as the concept of big data gradually gained traction, scholars began to focus on the issue of data pricing. Liu Zhaoyang [12] proposed a pricing model for big data from a mathematical model perspective, outlining seven pricing strategies. He pointed out that only prices that can provide positive feedback can lead to growth in the scale of data transactions, and it is necessary to make reasonable use of negotiation mechanisms to promote transactions. Wang Wenping noted that different pricing strategies are suitable for different types of data, with common strategies including platform set pricing, fixed pricing, real-time pricing, negotiated pricing, and auction pricing [13]. In 2021, Lu Minfeng and others focused on analyzing research related to data valuation pricing models and creatively proposed a two-tier market mechanism for data [14]. The primary data market is used for the production and valuation of data, similar to setting a listing standard for data, while the secondary data market is used for circulation and market-based pricing. Guo Xinxin and others in 2022 proposed that unilateral pricing by data trading platforms can lead to unfairness, making it difficult for buyers to obtain reasonable pricing, and demonstrated the effectiveness of data auction pricing mechanisms through operations research models [15].

However, the above studies all focus on pricing existing data or data directly obtained by enterprises and governments in operation, essentially addressing how to facilitate data transactions and promote data listings. They do not lead to the formation of a new productive force in data elements.

B. Data Rights Confirmation

With the development of data transaction technology, the study of data ownership has caught the attention of many scholars, but corresponding regulations remain elusive. Due to the unclear production process of data itself, the initial ownership of data is hard to determine. It is also challenging to unify whether the transaction involves data ownership or the right to use the data, and how to handle cross-platform secondary sales of data, among other issues of data ownership, which restrict the legal, compliant, and reasonable circulation of data elements. Tian Jietang [16]summarizes these as the heterogeneity and non-exclusivity of data, and the multiplicity of rights holders, with the definition of data property rights currently being a significant controversy in both academic and industrial circles.

In scenarios where enterprises control data, if the metadata is accumulated from the usage of various users, there will be a conflict of interest between the data subjects and the data controllers. It does not align with the data element value theory under the Web3.0 perspective, which thrives on everyone's participation and shared benefits. Therefore, another classic faction believes that individual users' data property rights should not be excluded due to data privacy rights, and that individual data's initial property rights should be recognized even if the data does not contain personal privacy [17]. Shi Dan also suggests that enterprises should request authorization from individual users to transfer data usage rights to promote the development of new data products [18].

C. Data Trading

Data trading is the core content of data element circulation, akin to financial markets, where future data assets should be traded on the exchange floor [19]. For example, the Shanghai Data Exchange is actively exploring the path for data assets to enter the ledger, proposing a "three-step sublimation method" for data assetization: from data resources to data products, from data products to tradable data products, and from tradable data products to data assets, forming an observable and measurable path for data assetization based on data product transactions.

Fig. 2. Shanghai data exchange data trading process.

However, this traditional method of data trading places too much emphasis on compliance and listing, which raises the participation threshold for small-scale and individual data providers, reducing the subjective enthusiasm of data providers, and is not conducive to achieving universal participation in the co-creation and sharing of data elements. It also fails to enable data demanders to obtain data that does not yet exist but is needed, making it difficult to build a thriving ecosystem for the full process of data element value release.

Overall, data trading is driven by the demand for data rather than being determined by the supply side. Drawing from the development experience of foreign data markets, establishing unified and centralized data sharing platforms by governments can increase the enthusiasm for data collection and production, primarily relying on government endorsement and credibility [20]. For instance, websites like data.gov.xx in countries such as the United States, Singapore, and India serve as examples. However, the domain data.gov.cn in China has not been activated yet, indicating the absence of a unified government-led open data platform. Researchers abroad studying data markets generally advocate for the involvement of data intermediaries, or data brokers, in data circulation and transactions to promote data sharing [21]. Scholars like Bertin et al. have introduced the concept of data market failure, attributing it to negative externalities, transaction costs, and information asymmetry, which they believe can be addressed by nurturing capable groups of data brokers to unlock the potential value of data and match buyers effectively [22].

The overall layout of the data market, the ecological context, and the design of interconnectivity pathways can draw lessons from the latest proposals by international scholars.

In 2020, Yan Li and others emphasized the indispensable role of a trusted third party in data transactions and pointed out that the decentralized characteristics of blockchain can solve single-point failure issues and achieve fair data transactions [23]. The paper proposes two different models: using homomorphic encryption and data samples to improve system reliability and data availability for fair data transactions, and preventing signature tampering in combination with smart contracts through dual authentication to achieve fairness in data transactions. Experimental results demonstrate the practicality of these data transaction smart contracts; however, the paper does not discuss in detail how to initiate a cold start of the data market or propose a complete method for constructing a data market ecosystem.

In 2021, Donghui Hu and others further refined their blockchain-based big data trading system, DataTBC, originally proposed by their team the previous year [24]. This system provides data requester with the ability to post requirements via smart contracts and facilitates price negotiations between parties. It combines data encryption with smart contracts, where encrypted datasets can only be accessed upon successful execution of the contracts. Lastly, DataTBC utilizes its decentralized nature to allow data buyers to participate in data assessment, directly impacting data quality ratings and revenue distribution. This design and strategy effectively enhance the security, availability, and fairness of data transactions.

In 2022, Trent proposed the Ocean Protocol as a substantial advancement in the data trading market. It promotes the development of the Web3 data economy using data tokens (based on the ERC20 token standard) that represent the right to access data services [25]. The protocol aims to democratize data by supporting its publication, sale, and consumption. It leverages smart contracts and blockchain technology to ensure data privacy, interoperability, and secure transactions. An automated market maker (AMM) for price discovery and "compute-to-data" to monetize data while protecting privacy were introduced. Additionally, OceanDao, a decentralized data trade governance organization, was established to fund and assist with the development and promotion of data-related software.

The working mechanism of the Ocean Protocol involves constructing data token pools and integrating automated market makers (AMMs) to facilitate data transactions. Data owners can issue uniquely identified data tokens based on the Ocean Protocol. When data requester purchase (swap) these tokens using Ocean tokens, they gain the right to download the data, and the corresponding data token price will increase to enable automatic pricing. The right to access data is tied to the possession of data tokens, and unauthorized possession of data will be subject to legal liability for infringement. Ocean Protocol has built a broad, public ecosystem that enables secure data sharing and monetization, supporting a sustainably growing data economy.

The design of the Ocean Protocol is nearing perfection, restoring the initiative in data circulation back to the data producers. However, issuing data requires building one's own data token liquidity pool. If the price of Ocean tokens is too high, it could increase the economic burden on small and medium-sized data holders to deploy and share their data, potentially reducing the number of data listings.

Building on previous research, in Section 4, we will outline a model for a Web3.0 data element trading market that creates a positive feedback loop to unleash the value of data elements effectively.

IV. Value Release Strategy Analysis

Through the analysis of existing research on the data element market, we have identified that current studies generally focus more on the data processing technologies after data providers supply the data, while neglecting how to motivate data providers to participate more actively in the data element market. This oversight in compensating the providers for their value may, to some extent, lead to a lack of long-term motivation for providers to supply data, resulting in a mismatch between supply and demand in the data element market, insufficient market vitality, and unsustainable development. To address this, we introduce the concept of bounty production. Utilizing the transparency of blockchain smart contracts, we envisage a balanced data element ecological reference model aimed at enhancing production efficiency and resource allocation efficiency, maximizing the potential of data element value release, and achieving sustainable development of the data element market.

In the context of Web3.0, the value release of data elements is conceived as a complex and comprehensive system, intended to facilitate efficient data circulation, value appreciation, and fair trading through innovative technologies and models. This chapter will provide a detailed explanation of the overall circulation model, value circulation concepts, and the roles of various member organizations.

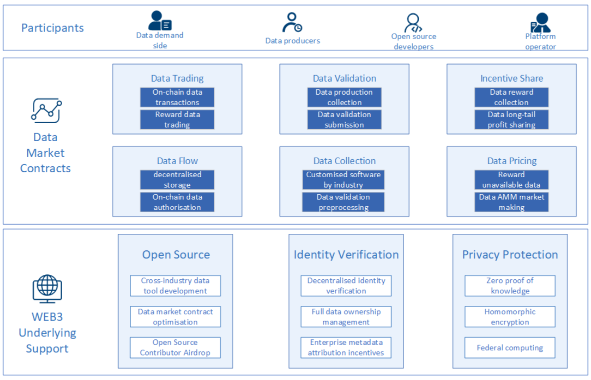

Fig. 3. Web3 data element market model.

A. Relevant Roles in the Data Element Market

1) Data Demand Side: In the Web3.0 data element ecosystem, the data demand side, or the buyers, play a crucial role. They facilitate the acquisition of data elements that meet specific business needs by clearly expressing their data requirements and providing the necessary financial support. Upon acquisition, the data elements are desensitized and stripped of personal information to ensure their safe application in a privacy-protected environment. The active participation of the data demand side not only injects capital into the market but also provides clear demand directions, forming the core force that drives the efficient operation of the market.

The data demand side not only stimulates the vitality of data production and supply but also drives the development and innovation of data services and solutions, thereby accelerating the maturity of the entire data ecosystem. Therefore, the data demand side is crucial for the healthy development of the data element market, playing an indispensable role in promoting data circulation, value-added service innovation, and maintaining dynamic market balance.

2) Web3 Intermediaries: In the Web3.0 data ecosystem, intermediaries create platforms that centralize various data-related services and tools, significantly reducing the friction costs of transactions among market participants. This platform effect not only facilitates transactions between buyers and sellers but also attracts more participants into the market, increasing its liquidity. As the number of participants grows, the variety and quality of data on the platform also improve, creating a positive feedback loop that further enhances the platform's attractiveness and competitiveness.

The presence of intermediaries significantly enhances the efficiency and activity of the data market. Without intermediaries, buyers and sellers might need to invest a lot of time and resources in finding counterparties, verifying data authenticity and quality, and managing transaction execution. Intermediaries, by providing one-stop services, greatly simplify this process, reduce transaction costs, and accelerate transaction speed, thus proving their necessity in the data market.

3) Data Producers: In the Web3.0 data ecosystem, data producers, also known as sellers, play a central role in supplying data. They produce and supply data based on market demands, encompassing a wide range of entities including but not limited to individual users, community members, businesses, and other organizations. By selling data on the market, sellers not only receive economic returns but also contribute to the prosperity of the entire ecosystem. In this process, particular emphasis is placed on protecting the sellers' privacy rights and data ownership, ensuring that their interests and rights are fully respected when sharing and trading data.

This role is crucial in driving the development of the data economy and promoting knowledge sharing. By meeting data needs from various fields, data producers not only increase the availability and diversity of data but also foster cross-disciplinary innovation. The decentralized nature of the Web3.0 ecosystem further enhances the transparency and fairness of data transactions, providing a more open and secure environment for data producers. This, in turn, motivates more participants to join the data production and supply chain, collectively advancing the continuous prosperity and development of the data ecosystem.

4) Open Source Contributors: Open source contributors play a central role in the strategy for releasing value from data elements in Web3.0. They facilitate the development of the entire ecosystem by developing and providing tools for data element production, participating in Web3 DAOs (Decentralized Autonomous Organizations), and governing intermediary components. The work of open source contributors is not limited to coding; it also includes designing reward mechanisms, reducing development costs, and formulating strategies to prevent malicious activities.

The adoption of the open-source model provides several key advantages to the entire ecosystem:

- Reduced Development Costs: By sharing resources and tools, duplication of effort is minimized, which enhances development efficiency and thereby reduces the overall development cost of the ecosystem.
- Open Code for Audit: Open source code can be reviewed by anyone, this transparency helps improve the quality and security of the software, and increases user trust in the system.
- Decentralization: Open source contributors work in a distributed manner, promoting the practice of decentralization, aligning with the core philosophy of Web3.0.
- Shared Prosperity: The open-source culture encourages the free sharing of knowledge and technology, promoting the rapid spread and application of innovation, which benefits the prosperity and sharing within the entire ecosystem.

Through these activities, open source contributors lay a solid foundation for building an open, transparent, secure, and efficient Web3.0 data element market ecosystem, fostering technological innovation and the healthy development of the ecosystem.

B. Data Element Market Circulation Model

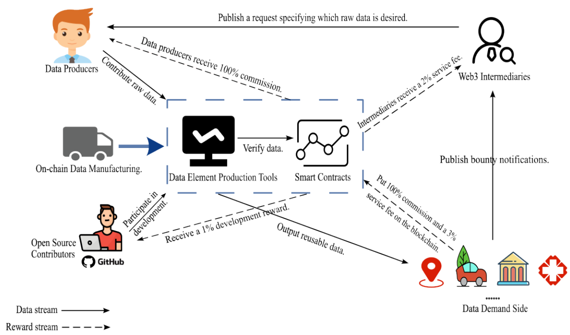

Fig. 4. Data element market participant roles diagram.

Within the Data Element Market Circulation Model, the data element market is designed to support the efficient circulation of data from producers to consumers, while ensuring data security and privacy. This model relies on blockchain technology to provide clear data ownership, transparent and secure transactions through decentralization and encryption. Additionally, the market circulation model includes technologies such as smart contracts and decentralized identity verification (DID), to automate transaction processes, ensure trust between parties, and reduce the cost and complexity of intermediaries.

Blockchain technology provides the infrastructure for data circulation, supporting core functionalities such as decentralization, encryption protection, and smart contracts. These technological innovations not only ensure the security and transparency of data circulation but also allow for the automatic execution of contract terms, thereby reducing manual intervention and associated costs, and enhancing transaction efficiency.

Under the DataR2E model in the data element market, data demand side initiate the market by publishing data requirements and offering incentives to attract data providers who hold the needed data or individuals willing to collect and produce data to participate in data verification and submission. As the amount of verified data in the market gradually increases, both the original data demand side and producers can capture the long-tail value created by the initial data, thereby motivating both buyers and sellers in this data element market to produce and purchase. Ultimately, this leads to the formation of a unified national or international largest data element market, realizing the universal sharing of data element value release.

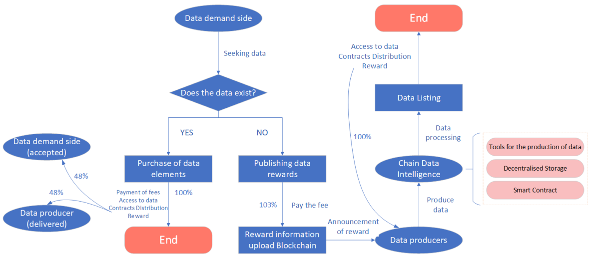

Fig. 5. Data acquisition process flowchart.

C. Data Element Production Tools

Data Element Production Tools refer to those tools and platforms that support data producers in quickly and efficiently producing data. These tools may include functions such as data collection, editing, validation, and processing, aimed at lowering the barriers to data production and increasing efficiency.

Within this framework, data demand side are not only users of the tools but can also act as open source contributors, jointly developing and refining open-source data collection tools with Web3.0 project platforms. Through this collaboration, data demand side can earn credits in the open-source community, which can offset future data acquisition costs, achieving a cost-effective win-win situation. For example, in the automotive data industry alliance, various car manufacturers can join together to develop standardized automotive data collection and production validation tools that not only promote standardization of data sharing within the industry but also accelerate the output of high-quality data.

This open development model not only deepens the understanding of the data production process and tool requirements among all parties but also promotes cross-enterprise and cross-industry collaboration, driving innovation and progress in data production and collection technologies. Moreover, this model aligns with the decentralized spirit of Web3.0, stimulating broader community participation through a decentralized contribution and reward system, thereby contributing to the prosperity of the data ecosystem.

D. Data Element Bounty Production Scheme

1) Pricing Scheme for Data Element Bounty Tasks: A crucial issue for maintaining the vitality of value circulation in the market is whether a reasonable pricing of data requests can be achieved when data demand side post bounties. This paper does not focus on detailed specific data pricing schemes but provides a macro, implementable market conceptual framework. More mature pricing and distribution schemes will be explored in future research. Currently, we can refer to the Non-Fungible Token (NFT) market mechanism, where if an NFT does not find a buyer in the trading market, the seller typically lowers the price until a transaction occurs. In the data bounty context, contracts can be designed to allow for the increase of commissions to stimulate more participants to complete bounty tasks.

2) Distribution Scheme for Data Element Bounty Tasks: Establishing appropriate thresholds for distributing data production tasks is crucial to ensuring an orderly market startup during the initial stages, avoiding an influx of data producers. As the market matures and expands, the number of data producers may surpass the actual demand. To control this potential trend effectively, our task allocation scheme involves rigorous qualification assessments before accepting tasks, selecting suitable data producers to ensure the correct execution of tasks, and preventing data collection failures due to full upload slots.

For candidates accepting bounty tasks, we plan to implement a priority ranking mechanism that could utilize log feedback from data production tools or adopt a federated learning cross-training method to calculate data value in a decentralized manner. This process could be implemented through platforms similar to Flock.io, helping to prioritize data producers.

Furthermore, if no threshold for task crowdsourcing is set, adopting a "first-come, first-served" (FCFS) model might lead to a decline in overall data quality. By entering metadata in data production tools, on-demand submission and real-time verification can be achieved, avoiding excessive data accumulation. Flexible adjustment of metadata bounty prices can effectively control the number of data providers, ensuring that the uploaded metadata meets entry standards and reducing malicious competition. Finally, introducing a data evaluation mechanism will help mitigate data quality issues caused by time competition.

3) Reward Distribution Scheme for Multiple Data Producers: With multiple, even numerous data producers in the trading market, fair distribution of data bounties among multiple data producers becomes a key issue. Our architectural philosophy strongly supports multiple data producers contributing simultaneously to a single data element demand, which can enhance the vitality of the data element market. Here are two envisioned bounty distribution schemes:

- First Come First Service (FCFS), cap reached: Bounties are first-come, first-served until the accepted data is full. Task distribution strictly refers to the submission time of the data producers, and bounty distribution is based on the quantity of metadata submitted successfully before the total data requirement is met.
- Deadline, contribution distribution: If the total collection amount is not completed within the specified time, bounties are distributed proportionally; if collected in excess within the specified time, data quality is first evaluated before assessing profit distribution.

Regarding bounty distribution, we might consider basing bounty distribution on the actual usage proportion of the provided data elements demanded, but more mature distribution schemes will be refined as research progresses.

V. Conclusion

We introduced the concept of data elements and the formation of the data element market, and addresses issues such as privacy protection that arise during the data circulation process. It provides a detailed exploration of data integration, exchange sharing, privacy computing, and blockchain technology methods, and analyzes the current development of the data element market. In response to current challenges such as the low sustainability of the data element market and the lack of enthusiasm among data providers, we introduce the concept of data element bounty production. Utilizing the transparency of blockchain smart contracts, we envisage a value release reference model that promotes the sustainable development of the data element market, enhancing production efficiency and resource allocation efficiency, thereby maximizing the value release of data elements. Finally, we also present comparative thoughts and future perspectives on key detail issues concerning this market concept.

Acknowledgment

The authors would like to thank the National Key Research and Development Program of China (No. 2023YFB2703900).

References

[1] "White paper on the secure circulation of data elements.." http://www.caict.ac.cn/kxyj/qwfb/bps/202301/P020230107392254519512, 2022.

[2] Yuncheng Shen, Bing Guo, Yan Shen, Xuliang Duan, Xiangqian Dong, and Hong Zhang, "A pricing model for Big Personal Data," vol. 21, no. 5, pp. 482–490.

[3] Y. Zeng, "To unleash the value of data elements, it is crucial to safeguard the security baseline.," *State-owned enterprise management*, p. 31, 2023.

[4] Q. Liu, "Web3.0 facilitates the valorization of data elements," *China's Industry and Information Technology*, pp. 42–49, 2023.

[5] H. He, A. Yan, and Z. Chen, "Survey of smart contract technology and application based on blockchain," *Computer Research and Development*, vol. 55, pp. 2452–2466, 2018.

[6] Q. Yao, "Web3.0: the coming new internet," *Chinese Finance*, pp. 14–17, 2022.

[7] Hangzhou Hyperchain Technology Co., Ltd., "Trusted data circulation white paper.." https://www.hyperchain.cn/news/559.html, 2022.

[8] B. Shen, J. Guo, and Y. Yang, "Medchain: Efficient healthcare data sharing via blockchain," *Applied sciences*, vol. 9, no. 6, p. 1207, 2019.

[9] L. Zero One Think Tank Information Technology (Beijing) Co., L. Bai, and B. Yu, *Privacy Computing: New Infrastructure for the Digital Economy*. Beijing Book Co. Inc. Google-Books-ID: h9asEAAAQBAJ.

[10] S. Nakamoto and A. Bitcoin, "A peer-to-peer electronic cash system," *Bitcoin.–URL: https://bitcoin. org/bitcoin. pdf*, vol. 4, no. 2, p. 15, 2008.

[11] Z. Liu, Y. Xiang, J. Shi, P. Gao, H. Wang, X. Xiao, B. Wen, Q. Li, and Y.-C. Hu, "Make web3. 0 connected," *IEEE transactions on dependable and secure computing*, vol. 19, no. 5, pp. 2965–2981, 2021.

[12] C. Liu, "Analysis of big data pricing issues," *Library and Information Science Knowledge*, pp. 57–64, 2016.

[13] W. Wang, "Research on pricing strategies for big data transactions," *Software*, vol. 37, pp. 94–97, 2016.

[14] M. Lu and W. Ouyang, "Research on the marketization of data elements and the institutional mechanisms for data asset valuation and pricing," *Xinjiang Social Science*, pp. 43–53+168, 2021.

[15] X. Guo, Q. Li, H. Wang, and J. Du, "Research on data transaction auction pricing mechanisms under information asymmetry in demand," *Operations Research and Management*, vol. 32, no. 11, pp. 170–175, 2023.

[16] J. Tian and L. Liu, "Trading models, rights definition, and cultivation of data element markets," *Reform*, no. 07, pp. 17–26, 2020.

[17] D. Xiao and Y. Wen, "Outline of data rights genealogy," *Journal of Xiangtan University (Philosophy and Social Sciences Edition)*, vol. 39, pp. 69–75, 2015.

[18] D. shi, "Research on data ownership and its protection approaches in the era of big data," *Journal of Xi'an Jiaotong University (Social Sciences Edition)*, vol. 38, pp. 78–85, 2018.

[19] H. Hou, "Exploration and practice of data trading in the context of factorization," *Industrial Innovation Research*, pp. 60–62, 2023.

[20] Y. Dou, M. Guo, L. Zhang, Z. Sun, and L. Zeng, "Research on the overall layout and promotion path of the national integrated data trading venue system," *E-Government*, pp. 2–11, 2024.

[21] F. Commission, "Data brokers: A call for transparency and accountability," pp. 1–101.

[22] B. Martens, A. De Streel, I. Graef, T. Tombal, and N. Duch-Brown, "Business-to-business data sharing: An economic and legal analysis," *EU Science Hub*, 2020.

[23] Y. Li, L. Li, Y. Zhao, N. Guizani, Y. Yu, and X. Du, "Toward decentralized fair data trading based on blockchain," *IEEE Network*, vol. 35, no. 1, pp. 304–310, 2020.

[24] D. Hu, Y. Li, L. Pan, M. Li, and S. Zheng, "A blockchain-based trading system for big data," *Computer Networks*, vol. 191, p. 107994, 2021.

[25] N. Dong, Z. Wang, J. Sun, M. Kampffmeyer, W. Knottenbelt, and E. Xing, "Defending against poisoning attacks in federated learning with blockchain," *IEEE Transactions on Artificial Intelligence*, 2024.