Blockchain for Data Marketplace: Enhancing Security, Privacy, and Trust

Silvio Meneguzzo (PhD in Blockchain and DLT, University of Camerino, University of Turin), Alfredo Favenza (Links Foundation), Valentina Gatteschi (Politecnico di Torino), Claudio Schifanella (University of Turin)

***Abstract*—In this oral communication, we aim to propose the concept of an energy data marketplace and the challenges associated with it, such as data quality, trust, privacy, and security. We will discuss how blockchain and DLT can provide transparency, immutability, and decentralization to enhance the security, privacy, and trust of data marketplaces. We will also present Data Cellar, which is a conceptual architecture for a blockchain-based data marketplace and discuss its key components. Furthermore, we will demonstrate the feasibility and effectiveness of our proposed approach through a specific case study. Finally, we will highlight the potential of blockchain and DLT in the data marketplace context and their impact on the future of data exchange.**

I. Introduction

Data marketplaces have emerged as a promising solution to the increasing demand for data-driven decision-making in various domains, including energy, healthcare, finance, and marketing [1]. They enable data owners to monetize their data by making it available for sale to interested parties, such as researchers, companies, and governments. However, the adoption of data marketplaces is hindered by several challenges, such as data quality, trust, privacy, and security. These challenges are particularly relevant in the context of sensitive data, such as personal and medical data.

To address these challenges, blockchain and DLT have been proposed as key infrastructures for building data marketplaces [2]. They provide transparency, immutability, and decentralization, which enhance the security, privacy, and trust of data marketplaces. In this oral communication, we will discuss the advantages of blockchain and DLT in addressing the challenges associated with data marketplaces and present a conceptual architecture for a blockchain-based data marketplace in the energy sector. We will showcase the feasibility and effectiveness of our proposed approach through a case study.

The practical implications of blockchain-based data marketplaces could be significant for businesses and policymakers looking to implement secure and transparent data exchange and for breaking data silos.

II. Characteristics of data marketplaces

A data marketplace is a platform that enables the exchange of data between data providers and data consumers. It provides a single point of access where data owners can offer their data for sale, and data consumers can search for and purchase the data they need. Data marketplaces can offer various types of data, such as public, proprietary, etc., depending on the marketplace's purpose and target audience.

Data marketplaces furnish several benefits to various stakeholders, including data providers, data consumers, and society as a whole. They can lead to increased revenue and new business opportunities for data owners by enabling them to monetize their data. Data consumers can access a wide range of data sources that can be useful for research, product development, and decision-making. Data marketplaces can ensure data quality by verifying and validating the data offered by data providers, which can lead to more reliable and accurate results. They can also foster collaboration and innovation by bringing together data providers and data consumers from different domains, leading to new insights and discoveries.

III. Use case and how blockchain is integrated

A greener energy system is crucial to the future prosperity and liveability of European citizens. It is a transformation that is cultural, psychological, and economical. Above all, it creates a vast amount of new data: new load profiles, the rise of prosumers, the uptake of electric vehicles, the addition of storage. This data avalanche is very well expressed in the European policy directives i.e., the European Green Deal proposal [3], the Industrial Strategy, the Circular Economy Action Plan [4], and the Digital Strategy, already recognize that clean energy transition and digitalization go hand in hand. The EU has defined a Blockchain strategy [5] striving to be a leader in the field, acting as an innovator in blockchain housing significant platforms, applications, and companies. The base of the strategies is to pursue several "Gold Standards" including 1) Environmental Sustainability, 2) Data Protection, 3) Digital Identity, 4) Cybersecurity, and 5) interoperability. Prompted by the previous rational and background drivers, supported by the opportunities, and following the European strategies, Data Cellar is our use case that will develop a dynamic, interoperable energy-oriented data platform to support the uptake of the Energy Communities leveraging a blockchain-based tokenization scheme for the remuneration in data and pre-trained AI models provisioning/acquisition cycle [6].

Data cellar is the marketplace where different entities will be able to trade (buy, rent and sell) data and pre-trained AI models, leveraging algorithms and software libraries for processing the collected data and simplify operations of data analytics and business intelligence thus aligning the incentives of the stakeholders through an on-chain tokenization scheme. The final objective is to provide a trusted environment to all involved participants: data providers, service providers, and service consumers.

IV. Data Cellar conceptual Architecture

The architecture of Data Cellar Marketplace, a blockchain-based data marketplace, is based on the Ocean Protocol

[2]. We selected Ocean Protocol as the tech stack for our use case due to its advanced features for data monetization, access control, and governance. We also customized the Ocean Protocol architecture to meet the requirements of Data Cellar.

A. Architecture Components

In Fig. 1 the conceptual architecture of the Data Cellar use case is presented.

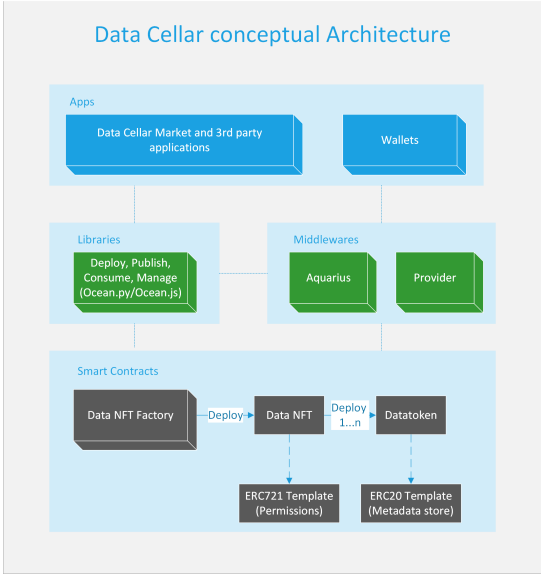

Fig. 1. Data Cellar Architecture Overview

1) Apps Layer: The Applications layer is the topmost layer of our architecture. This layer in our use case consists of applications such as Data Cellar Market that provide users with a platform to onboard services like data, algorithms, and compute-to-data into crypto. Users can publish and mint data NFTs and datatokens, hold datatokens as assets in data wallets, discover assets, buy/sell datatokens for a fixed or auto-determined price on data marketplaces, and use data services by spending datatokens.

2) Libraries Layer: The Libraries layer consists of libraries used by the applications to interact with the smart contracts. These libraries encapsulate the calls to the smart contracts and provide features like publishing new assets, facilitating consumption and much more. Additionally, the libraries also integrate with Provider to provision and access data services and Ocean Aquarius for metadata.

3) Middleware Layer: The Middleware layer provides the necessary infrastructure for discovering and accessing data assets on Ocean Protocol. This layer comprises Aquarius, a metadata cache for faster search, and Provider, which facilitates downloading assets, DDO encryption, and communicating with operator-service for Compute-to-Data jobs.

4) Smart Contracts Layer: The Smart Contracts layer is the lowest layer of the our architecture, and it is responsible for managing the interactions between data providers, consumers, and the network. This layer consists of smart contracts deployed on the Ethereum mainnet. These smart contracts are used to facilitate the creation, management, and trade of datatokens. The libraries interact with these smart contracts to provide features like publishing new assets, facilitating consumption, and much more. Additionally, the smart contracts support various token standards like ERC721 and ERC20 to represent data assets and licenses to access them. The smart contracts also support Access Control Tools that enable publishers to manage access to their data assets.

B. Addressing Challenges

Blockchain-based data marketplaces can address the challenges associated with traditional data marketplaces, such as data quality, trust, privacy, and security.

1) Data Quality: Blockchain-based data marketplaces can ensure the quality and accuracy of data by implementing a reputation system that rates data providers based on the quality and accuracy of their data. Additionally, smart contracts can be used to enforce data quality standards and automatically validate and verify the data.

2) Trust: Blockchain-based data marketplaces can enhance trust between data providers and data consumers by providing transparency, immutability, and decentralization. All transactions are recorded on the blockchain, which can be audited and verified by all participants. Smart contracts can also enforce data usage policies and handle payments, reducing the risk of fraud and ensuring fair transactions.

3) Privacy: Blockchain-based data marketplaces can ensure privacy by implementing privacy-preserving techniques, such as encryption and pseudonymization. Data providers can choose to anonymize their data or provide access to only certain parts of their data. Additionally, data consumers can be required to provide proof of identity and a valid reason for accessing the data.

4) Security: Blockchain-based data marketplaces can enhance security by leveraging the security features of the blockchain network, such as encryption, cryptographic hashing, and consensus mechanisms. Smart contracts can also ensure secure transactions by handling payments and enforcing data usage policies.

V. Conclusion

In this oral communication, we presented the concept of data marketplaces and their associated challenges, such as data quality, trust, privacy, and security. We proposed a blockchain-based approach to building an energy data marketplace (Data Cellar) and, highlighted the advantages of using blockchain and distributed ledger technology to address these challenges.

References

[1] Amila Saputhanthri, Chamitha De Alwis, and Madhusanka Liyanage. Survey on blockchain-based iot payment and marketplaces. *IEEE Access*, 10:103411–103437, 2022.

[2] Ocean Protocol Foundation Ltd. Data: The new asset class. https://oceanprotocol.com/,.

[3] The European Commission. A european green deal. https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/european-green-deal$_en$, .

[4] The European Commission. Circular economy action plan. https://environment.ec.europa.eu/strategy/circular-economy-action-plan$_en$, .

[5] The European Commission. Blockchain strategy. https://digital-strategy.ec.europa.eu/en/policies/blockchain-strategy,.

[6] European Union under the Horizon Europe Framework Programme, Grant Agreement No. 101069694. Data cellar. https://datacellarproject.eu/,.