DataSail: Creation of a User Data Market System Through the Cardano Network

***Abstract*—In an era where data is often described as the new oil, control and monetization of personal data have largely remained in the hands of centralized entities. Leveraging the power of smart contracts on the Cardano blockchain, a decentralized data marketplace is introduced, empowering individuals to securely tokenize and trade their personal data. This approach ensures privacy, autonomy, and direct compensation for users, providing a transparent framework for a more equitable data economy.**

I. Introduction

In the modern era, data has become one of the most valuable assets, playing a fundamental role in shaping our economies and lives. It drives decision-making in various sectors, from healthcare and finance to marketing and public policy. However, the advent of blockchain technology offers a potential into a future where digital asset management is more secure and decentralized. Despite these advancements, the majority of data is still controlled by a few large organizations, raising concerns about privacy and the autonomy of individuals. This centralization results in a significant disparity: individuals, the primary creators of data, receive minimal benefits from sharing their information, often only in the form of slightly improved services or personal experience.

Moreover, the lack of proper rewards or control over their own data leaves many unaware of its potential value. This issue is becoming increasingly critical as the demand for data, driven by developments in machine learning and artificial intelligence, continues to rise.

Addressing this challenge, this paper introduces a novel decentralized data marketplace developed on the Cardano blockchain, aiming to facilitate automatic data transactions between buyers and sellers. By advocating towards a system that is both fair and transparent, it can empower individuals with control and direct benefit over their personal data.

Through the usage of Cardano's Extended UTxO [1] model and smart contracts, the marketplace enables users to tokenize their data, thereby converting it into tradable assets. DataSail incorporates a browser extension for data collection, IPFS [2] for decentralized data storage, and cryptographic techniques for secure data transmission and identity verification. During DataSail's demonstration, the attendees will have the chance to interact with the marketplace through an intuitive and easy-to-use UI, which will allow them to act both as data providers, being able to monetize their browsing data, and as data consumers, purchasing data for analytics purposes.

II. System Design

The architecture of DataSail [3] is an orchestration of several components: a browser extension, a dApp with front-end and back-end parts, and three Plutus smart contracts– two validation scripts and one minting policy. Additionally the InterPlanetary File System (IPFS) is employed for data storage, and a key-value store is utilized by the dApp.

Broadly speaking, DataSail accommodates two types of actors: Sellers and Buyers. A Seller is an individual who aims to monetize their browsing and personal data by selling it. On the other hand, a Buyer is an entity or individual interested in purchasing such data for analytics or other purposes. In the architecture of this decentralized marketplace, there are two primary transactional mechanisms that facilitate the exchange between a seller and a buyer: the *Ask* and the *Bid* flows.

In the *Ask* flow, the seller takes the initiative by locking a specific token [4] under a smart contract. The token is essentially "listed" with a predetermined price tag inside its Datum. This sets the stage for buyers to meet this price in order to unlock and acquire the token. The DataListing validator smart contract ensures that the token is securely held until the asking price is met, at which point the token is transferred to the buyer and the agreed-upon sum is sent to the seller.

Contrary to the *Ask* flow, the *Bid* model is buyer-centric. Here, buyers can place bids on tokens they are interested in. The bid consists of a certain amount of ADA that the buyer is willing to pay for the token. Sellers can browse these bids made on their tokens and choose to accept any that meet their valuation of the token. Upon acceptance, the smart contract facilitates the immediate exchange of the token and the bid amount between the seller and the buyer.

At the core of DataSail's functionality are three Plutus smart contracts, each serving a distinct purpose within the data marketplace ecosystem:

1. *DataToken Minting Policy*: This contract governs the creation of DataTokens, which represent the user's browsing data. Utilizing a parameterized minting policy, it ensures that DataTokens are uniquely minted by verifying the consumption of a specific UTxO. This policy prevents unauthorized token creation and enforces the singularity of each DataToken.
2. *DataListing Contract*: Operating under the *Ask* model, this contract enables sellers to list DataTokens for sale. It incorporates a validation mechanism to ensure that transactions meet the seller's specified conditions, including payment verification and the unlocking of only the intended DataToken, thus safeguarding against unauthorized access and ensuring transactional integrity.

3. *Bid Contract*: Supporting the *Bid* model, this contract allows buyers to place bids on DataTokens. It enhances discoverability and transaction security by generating unique addresses for each token and enforcing constraints to prevent double spending and ensure that bids are placed on unique tokens.

The back-end system, developed using Node.js and TypeScript, interfaces with the Cardano blockchain and IPFS (InterPlanetary File System) for data management. User browsing data, captured through a dedicated browser extension, is encrypted and stored on IPFS, ensuring privacy and security. Upon successful storage, the IPFS returns a Content Identifier (CID) unique to the encrypted data file. This CID is then embedded as metadata within the DataToken itself when minted on the blockchain. The use of Blockfrost as an IPFS gateway and a blockchain provider simplifies interactions with the Cardano blockchain and IPFS, streamlining the processes of data storage, retrieval, and transaction execution.

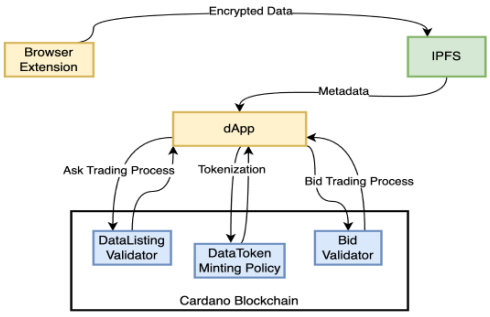

Fig. 1: High level Architecture diagram

Authorization within DataSail employs digital signatures and asymmetric cryptography to verify the identity of participants and secure transactions. This approach ensures that only legitimate owners can execute actions like minting, listing, or bidding on DataTokens. Specific mechanisms include:

- Digital Signature Verification: Utilized both client-side and server-side to authenticate user actions, ensuring that requests to mint or trade DataTokens are authorized and tamper-proof.
- Token Ownership Verification: Before enabling data retrieval or token trading, DataSail verifies token ownership through the blockchain, ensuring that only rightful owners can access or transact with DataTokens.

The design prioritizes security and privacy through several layers. First, browsing data is encrypted before storage on the IPFS, ensuring that sensitive information remains private and secure. Each smart contract includes security measures to validate transaction conditions, prevent unauthorized access, and ensure the integrity of the trading process. Finally, by leveraging blockchain technology and decentralized storage, DataSail minimizes reliance on centralized entities, reducing vulnerability to data breaches and enhancing user control over their data.

III. Demonstration Scenarios

This section demonstrates the operational capabilities of DataSail, focusing on the trading and management of browsing data as token on the Cardano blockchain. We illustrate the process from the perspectives of both the data seller (the user generating and selling browsing data) and the data buyer (the user interested in purchasing data tokens).

The demonstration is conducted on the Cardano preview to simulate real-world transactions without incurring real costs. DataSail communicates with the blockchain network through the lucid-cardano library and the Blockfrost provider.

A. Seller Experience

The seller activates the browser extension, inputs their wallet address, and selects the duration for data capture. Upon confirmation, a server receives the browsing history and encrypts it. The encrypted data is then stored on IPFS, and the result CID is associated with the seller's wallet address.

The seller then enters DataSail, which retrieves any data associated with their wallet. The dApp, through its UI, guides the seller in minting a DataToken representing their stored browsing data. The minting process leverages the DataToken Minting Policy smart contract. Once minted, a token listing is created and Bids can be placed on that Token (Bid flow). Otherwise, the seller could list the DataToken for sale using the DataListing smart contract, setting an asking price in ADA (Ask flow).

B. Buyer Experience

The buyer explores available DataTokens through DataSail's UI. Upon selecting an available DataToken of interest, the buyer places a bid on it using the Bid smart contract. This process locks the bid amount in ADA under the smart contract until the seller accepts the bid or the buyer retracts it.

For tokens listed directly for sale under the DataListing, the buyer can purchase them at the asking price. Once a transaction is completed (either through direct purchase or bid acceptance), the buyer can request the browsing data associated with their newly acquired DataToken. DataSail facilitates the decryption and download of the browsing data, ensuring that only the buyer, now the owner of the DataToken, can access the data.

All transactions, whether minting, listing, bidding, or purchasing, are confirmed on the Cardano blockchain, ensuring transparency and security. DataSail's backend and UI provide real-time feedback on transaction status, including confirmations and any errors.

References

[1] M. M. T. Chakravarty, J. Chapman, K. MacKenzie, O. Melkonian, M. P. Jones, and P. Wadler, "The extended utxo model," 2020.
[2] J. Benet, "Ipfs - content addressed, versioned, p2p file system," 2014, original IPFS white paper.
[3] "Data sail: Source code for decentralized data marketplace on the cardano blockchain," https://github.com/varagos/data-sail.
[4] M. M. T. Chakravarty, J. Chapman, K. MacKenzie, O. Melkonian, J. Müller, M. P. Jones, P. Vinogradova, and P. Wadler, "Native custom tokens in the extended utxo model," 2020.