DataMatch - The Future of Data Transactions

UC San Diego™
HALICIOĞLU DATA SCIENCE INSTITUTE

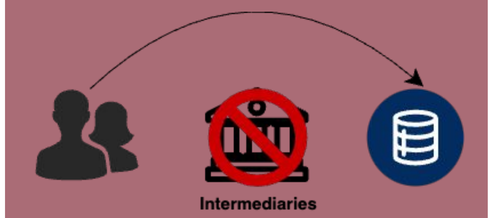
Aditya Melkote , Maxwell Fang, Yanhao Guo, Zixin Wei, Mentor: Sheffield Nolan
avmelkote@ucsd.edu, m3fang@ucsd.edu, yag007@ucsd.edu,
z5wei@ucsd.edu, Sheffield.Nolan@franklintempleton.com

Overview

DataMatch is an Ethereum-based data marketplace that allows anyone to buy or sell data securely *without* any intermediaries.

The platform leverages:
- Blockchain Technology
- PII Removal
- Off-chain storage solutions

Introduction



Problem
Current online data marketplaces:
- Reveal private information
- Require a hefty fees
- Contain scalability issues
- Centralized with single-point node failures

Solution
1. Smart Contracts
   a. **Immutable** and **decentralized**
2. PII Removal
   a. Removes **sensitive** information
3. InterPlanetary File System (IPFS)
   a. **Off-chain, decentralized** cloud storage

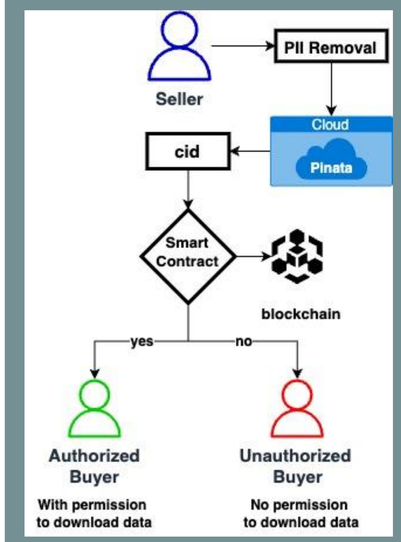
Background

What is blockchain?
- Blockchain is a public database that is updated and shared across many computers in a network.
- Decentralized -> no middle-man or authority
- Immutable -> blocks cannot be altered

What is Ethereum?
- Ethereum is a platform with blockchain technology
- Ether (ETH) is the native cryptocurrency
- Foundational for building apps and organizations in a decentralized and permissionless way, especially with smart contracts

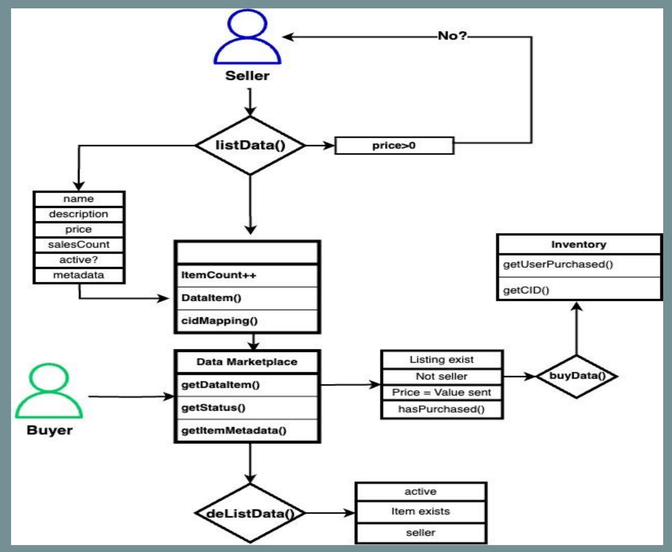
What are Smart Contracts?
- Vending machine: a script that performs actions if certain actions are satisfied e.g. receiving a dataset after purchasing it with crypto currency

Methodology

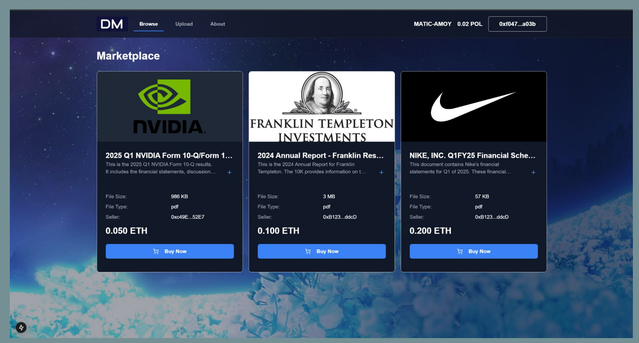


1. Seller uploads
2. PII checks for sensitive information
3. File uploaded to IPFS via Pinata
4. CID is generated and retrieved from IPFS
5. Smart contract stores listing on blockchain
6. Buyer purchases listing to gain permission to download data

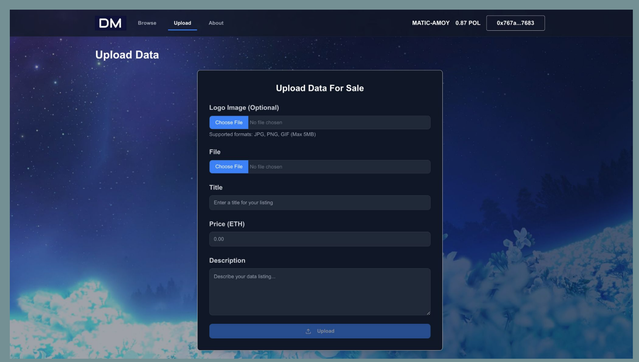
Smart Contract Flow-chart



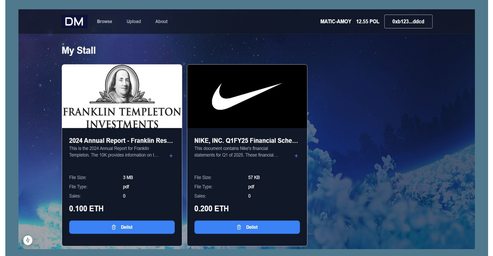
Browse



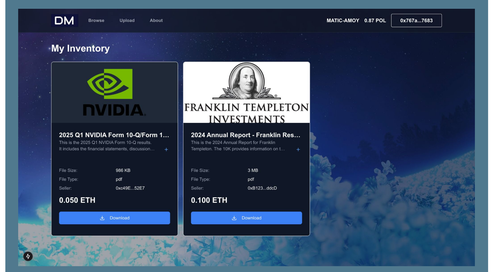
Upload



My Stall



Inventory



Future Work

- Improve the market interface to enhance user experience
- Implement a recommendation system to help user find desired dataset more efficiently
- Incorporate encryption techniques to prevent illegal usage of data

Acknowledgements

We would like to thank our mentor Sheffield, Professor Bellur, and TA Aritra Das for all the advice along the way.

Link to Website