A Framework for a Blockchain-Based Decentralized Data Marketplace

Meshari Aljohani(✉), Ravi Mukkamala, and Stephan Olariu

Department of Computer Science, Old Dominion University, Norfolk, VA 23529, USA
{maljo001,rmukkamag,solariu}@odu.edu

Abstract. Data is critical for improving an individual's quality of life. Its value provides opportunities for users to profit from data sales and purchases. Data marketplace users, on the other hand, must share and trade data in a secure and trusted environment while maintaining their privacy. The paper's first major contribution is to identify enabling technologies and challenges to the development of decentralized data marketplaces. The second major contribution is the proposal of a blockchain-based decentralized data marketplace framework. The proposed framework allows sellers and buyers to transact with greater confidence. The system employs a novel approach to enforcing honesty in data exchange among anonymous individuals by requiring a security deposit. The system has a time frame before the transaction is considered complete.

Keywords: Blockchain · Data · Data marketplace · Smart contract · Reputation

1 Introduction

Data is considered one of the most valuable assets for modern businesses. Companies like Google and Facebook provide free services in return for users' data [18]. So, the importance of data creates opportunities for users to make profits from selling and buying data. As a result, data marketplace is being the conventional framework that enables participants to trade their data. In a centralized marketplace, all transactions are directed through one central exchange. While data marketplaces enable individuals and companies to monetize their data, centralized data marketplaces suffer from several shortcomings. Single points of failure, data privacy, and security are examples of these shortcomings. Blockchains offer a potential answer to such problems. They help build a decentralized data marketplace. A decentralized data marketplace is one in which there is no central authority or server and where data owners retain control of their data and store it on their own devices [19].

Dealing with unknown sellers without a trusted third party (e.g., Amazon, eBay) could lead to cheating [11]. A reputation system is a solution for this kind of issue, allowing buyers and sellers to minimize their risks by selecting carefully

Z. J. Haas et al. (Eds.): WiCON 2022, LNICST 464, pp. 59–75, 2023.
https://doi.org/10.1007/978-3-031-27041-3_5

their transaction partners. Nevertheless, sellers with bad reputations try to avoid this by implementing different methods to make buyers believe they are good sellers [3].

Enforcing terms and conditions is challenging in a decentralized marketplace because there is no trusted authority or intermediary who could intervene into event of a dispute. So, once the transaction is submitted, there is no way for the buyer to get a refund. However, users are required to accept the terms and conditions of centralized marketplaces such as Amazon and eBay. Buyers can file a claim and request for a refund. Issues such as these have motivated us to propose a novel solution based on blockchains.

1.1 Contributions

In this paper, a data marketplace framework is proposed based on a novel idea of using a Dispute Resolution Center (DRC) and enforcing security deposits using smart contracts based on blockchains such as the Ethereum Blockchain.

The contribution includes the following:

1. Identifying challenges ahead;
2. Identify enabling technologies for decentralized data marketplaces;
3. Providing a framework for a decentralized data marketplace;

This paper is organized as follows: Sect. 2 summarizes the related work. Data Marketplace challenges and enabling technology are identified in Sect. 3 and 4, respectively. In Sect. 5, use cases are explained. In Sect. 6, the proposed Decentralized Data Marketplace Framework is described. In Sect. 7, the Global Data Marketplace is clarified. Finally, Sect. 8 summarizes and concludes the paper.

2 Related Work

In this section, various related work that used blockchain in data marketplace is presented. The main motivation of implementing blockchain in data marketplace is to provide a trust worthy environment while preserving the users privacy. Trust plays a significant role in marketplaces. Buyers are concerned when dealing with unknown sellers which affects their decisions on purchasing from that market [8]. Distributed Ledger Technology (DLT) was employed by [16] to establish a decentralized anonymous marketplace with a Sybil-attack resistant design. The system was implemented by utilizing existing blockchains and anonymous payment systems, such as Zerocash. Different approach was introduced by [6] to demonstrated a decentralized marketplace for private data. The system uses a smart contract for privacy-preserving of data, and enables data providers to impose their term of use, or constrain the use on their data. However, the system lacks a reputation system as well as a method for the buyer to dispute a transaction. Similarly, [21] proposed a solution that provides a reliable search and purchase method using Name Data Networking (NDN). However, the proposed

system does not include a dispute system. The authors in [15] offered a semi-decentralized data marketplace that provides micro services. Payment choices, storage options, and a metadata generator are available through the proposed system, which is implemented on local servers. For trading and valuing transactions, the smart contract-based blockchain is used.

Moreover, [19] proposed a decentralized data marketplace with three participants: sellers, buyers, and notaries. The buyer initiates the transaction by creating a data order and forwarding it to notaries to receive their fees and terms and conditions. The buyer then adds the list of notaries to the data order, together with their fees, terms of service, and signatures, and submits it to the smart contract. Sellers receive the data order, which contains the requested data, the suggested notary (just one), data price, and the terms and conditions of data use. The data response is then sent to the seller. When the buyer receives the sellers' data response, she selects a group of sellers and adds them to the contract. The data is then uploaded by the selected sellers to the specified address, which was sent by the buyer. The transaction is completed, and the payments are delivered after the buyer has the personal information. Nevertheless, the data buyer should verify with notaries to obtain their signed certificates before completing the data order. The limitation of this mechanism is that the notaries are paid whether the data is notarized or not, which adds an additional cost to the transaction.

3 Data Marketplace Challenges

Existing data marketplace frameworks are provided in the form of a centralized or decentralized systems. Centralized data marketplaces have many challenges that are addressed in decentralized data marketplaces. In this section, we present the most common challenges in centralized marketplaces.

3.1 Availability

Data availability refers to the ability to access a system's data when it is needed. Data marketplace systems store the required data to provide services for their users. Hence, it is important to ensure their availability to provide the needed data once it is requested. System updates, network breakdowns, and unpredictable faults are events prevent data from being accessed in a timely manner. More importantly, centralized data marketplaces are vulnerable to Single Point of Failure SPOF which is a feature in centralized systems that reduces its availability. SPOF means that the entire system becomes unavailable if the central node fails. Decentralized marketplaces, on the other hand, are designed to avoid the SPOF to provides higher availability since the other nodes are still available to access even if one of the other nodes fail.

3.2 Scalability

Scalability is one of the a main features in marketplaces to ensure that a marketplace is able to accommodate a large number of users. Centralized marketplaces, however, as any centralized system only scale at the server level, implying that only the central server scalability determines the number of users. Decentralized marketplace, on the other hand, incorporate additional number of servers in which each of them can serve additional number of users.

3.3 Privacy

The operators of centralized marketplaces may compromise the privacy of their users by knowing their real identities. Users on-site activities such as selling, buying, preferences, searching, etc. may be stored with their real identities and used for many purposes without user's knowledge. They also may reveal users' identities and personal information to a third party. In contrary, decentralized marketplaces identify their users with public keys that makes the users identity is anonymous. Hence, users data and information are not related to their real identities.

3.4 Security

Security is one of the most important requirements in data marketplace transactions. In centralized marketplaces, users share personal information such as real names, addresses, and sensitive data such as bank accounts and credit cards. A data breach can expose this information to malicious sources, putting them at risk. Moreover, data stored on the server could be stolen by attackers. On top of that, a single point of failure is one of the security issues for such systems, which are often exposed to malicious attempts such as DDoS attacks. Decentralized marketplaces, on the contrary, do not require users' information. Users can send and receive money under anonymous identities using public and private keys. The only concern is if the users share or lose their private keys.

4 Enabling Technologies

The purpose of this section is to highlight existing technologies that assist data marketplaces in overcoming the aforementioned issues.

4.1 Blockchain Technology

Blockchain technology is one of the important aspects that empowers marketplaces. It helps eliminate middlemen who are operating the system by implementing a decentralized system. Multiple decentralized nodes verify transactions on a blockchain, and anyone can join or leave the network at any time without affecting the network's capacity to achieve a consensus on transactions. Decentralization is a key to the technology that enables peer-to-peer exchange on the

Blockchain. It is a distributed, immutable ledger that supports the process of recording and tracking transactions. It is formed of connected blocks that are linked securely by using hard mathematical calculation and cryptographic hash function. Each participant has a private key and a public key, so any transaction is encrypted by using the public key and only decrypted by the accompanying private key [18].

In addition, data on the blockchain is immutable, so it can't be deleted or changed because it needs validation by the network. Also, marketplaces based on blockchain are more tolerant to the Single Point of Failure SPOF since data are distributed among the network nodes. They do not rely on a single entity; every member on the network has a copy of the ledger. Therefore, the risks of attacks like DDoS attacks can be avoided. Moreover, marketplaces based on the blockchain do not require intermediaries; the smart contract can be executed automatically. Also, payments do not need a third party, and they are instant.

IOTA is another blockchain technology that is a cryptocurrency and open-source distributed ledger built for the Internet of Things (IoT). It stores transactions in its ledger using a directed acyclic graph, which has the potential to be more scalable than blockchain-based distributed ledgers. The scalability of this system is achieved by removing the need for miners to confirm transactions and replacing them with nodes that want to add a new transaction but must first validate two previous transactions [16]. Therefore, transactions requiring micro-payments can be done easily and without cost [16]. The network obtains consensus through the IOTA Foundation's coordinator node which means it is currently centralized. Coordicide-IOTA 2.0, a new follow-up update, aims to eliminate the need for the coordinator for consensus [12]. IOTA has a number of benefits over Blockchains, including fee-free transactions, micro-transactions, and energy efficiency [10].

4.2 Artificial Intelligence (AI)

With the improvement of machine processing and the growth of data, artificial intelligence (AI) (computer vision, natural language, virtual assistants, robotic process automation, and advanced machine learning) has come to the surface. 70% of companies might have implemented at least one type of AI technology, but less than half will adopt all five types by 2030 as anticipated by [4].

Many researchers have proposed solutions related to reputation systems on marketplaces using machine learning. [7] provided a solution for the "all positive reputation" problem in the marketplace reputation system, where the average reputation score is 99 percent. They derive criteria based on opinion expressions using Natural Language Processing algorithms. They also present an approach that uses dependency relation analysis and the Latent Dirichlet Allocation (LDA) topic modeling technique to cluster these phrases into criteria and generate reputation rating and weight. Their research tries to eliminate positive bias in seller raking. The research was carried out in Hindi on eBay and TripAdvisor databases. LSTM Neural Network (NN) was used in [12] to analyze data col-

lected from a Facebook page of AliExpress, a well-known marketplace to classify positive and negative reviews.

Similarly, machine learning was used to predict users' reliability. [1] proposed a reputation model using from their profiles. They extracted users' ratings and feed them to the machine learning algorithms to calculate users' weight. Then they used it with a weighted average method to compute the final product quality score. They utilized multiple machine learning algorithms: linear regression (LR), support vector regression (SVR), K-nearest neighbor (KNN), and regression tree (RT).

5 Use Cases

In this section, various use cases are presented to demonstrate how the aforementioned enabling technologies alleviate data marketplace challenges. The use cases are categorised based on the data nature into two main types: static data use cases and real-time use cases to show how solutions tackle the challenges considering the data nature.

5.1 Static Data Use Cases

Static data refers to the fixed datasets or data that never changes after collecting. In data marketplace, static data is a one-time payment model. Customers purchase the static data as a single unit at once. The following two cases deal with static data.

Use Case A1: Healthcare Data

Consider the following scenario: a data marketplace selling healthcare datasets of diabetes patients. When a buyer wants to buy a dataset from the data marketplace, she looks for a seller with a good reputation and reasonable data pricing. The buyer negotiates the requirements with the seller. The seller wants to control her data by not uploading it to a third-party server. Instead, she will send it directly to the prospective buyer. As part of the negotiation, the seller provides metadata to the buyer, which is information about the data such as data description, number of records, file size, etc. Once they have reached an agreement, they both hash the offer and submit it to the smart contract, which acts as a middleman. In addition, the seller submits a security deposit, the URL of the encrypted data, and the decryption key, and the buyer submits payment and a security deposit.

Furthermore, this case has a privacy issue both parties desire to maintain their privacy. In this situation, technology like blockchain aids in the protection of the user's privacy and anonymity. Furthermore, because it compensates for the lack of authority, consumers can trade with confidence.

Use Case A2: Data Provenance

Many buyers are interested in purchasing data from well-known data providers. However, these providers usually don't sell data directly to buyers.

Instated, a third party collects data from these sources and resells it to buyers. For example a case as in `Use Case A1`, the buyer wants to buy diabetic patients data from a seller. The seller offers a dataset from Sentara Healthcare, which is a well-known healthcare provider. After the buyer received the data, she decides to dispute it and claim that the data was fake. Without relying on centralized authority, blockchain-based smart contracts serve as a trusted intermediary for data and payment exchange. Furthermore, artificial intelligence AI technologies such as deep learning and machine learning are possible solutions for such issues. In [9] and [17], authors leveraged AI technologies as solutions for data provenance.

5.2 Real-Time Data Use Cases

Real-time data is information that is available at the time it is generated. As soon as it is collected, the data is sent to users. In the interim, the data could be stored on the owner's server for later use. However, real-time data is a subscription model in which customers pay for services for some time. We provide two scenarios based on the data type. Real-time data is critical, and its value varies depending on the industry. Some businesses can tolerate a few seconds of delay, while others cannot. COVID-19 maps, which show the COVID-19 status for a specific location, are one example of this data.

It is difficult to exclude third parties when selling real-time data, especially when dealing with anonymous participants who want to maintain their anonymity. Furthermore, sellers want to be paid in full, whereas buyers require data for a specific time period. Users can meet these conditions with the help of a blockchain-based smart contract.

Use-Case B1: Weather Temperatures

Consider the case of a buyer looking to purchase real-time streaming data. If a buyer wishes to purchase weather temperatures from several locations, for example, she will choose a seller based on his reputation score and price. They then begin to negotiate over the requirements. The buyer, for example, specifies requirements such as acceptable latency, packet loss percentage, and service unavailability. They then submit the smart contract if they have reached an agreement as to the proposed framework. Until the contract expires, the seller begins transmitting data to the buyer.

If all goes well, the buyer will get his data and the seller will get his money. If, on the other hand, the buyer decides to be dishonest and claims that he did not receive the data for the time period specified in the contract, she will file a claim asking for a refund to a mediator, who is a person who works to reach an agreement amongst those who are involved in a dispute.

To eliminate the cost of engaging a third party and preserve privacy, blockchain is a solution where participants can ensure transact with each other safely. Also, it can track generated data by saving hash value as a footprint. So, the mediator, who is only needed when the buyer raises a dispute, can check on the recorded hash value. A smart contract can be implemented as proposed in section in this paper (Sect. 6) provides a trusted environment where both participants can exchange transactions without the involvement of the mediator.

Use-Case B2: Stock Market

Real-time data is extensively used in many trading applications, such as the stock market. Users can find different services in these applications, like stock quotes, last sales, best bid, and other services. Most of these services are free, but they do not provide real-time data. According to Reuters [13], all quotes are delayed by a minimum of 15 min.

Stock prices go up or go down second by second, which presents a fluctuation. That is why the current price is very important for users. It allows them to make decisions on price moves. On the contrary, trading on old data could cost a lot for users if they failed to detect an uptrend or downtrend within a stock.

In this case, a user is interested in real-time stock quotes. So, she subscribes to a service. She starts to use the service by selling, buying, or bidding based on the price. Since the service is subscription-based, if the user does not file a complaint by the end of the day, which means the transaction is completed and considered successful. If the trader files a complaint, it will go to a mediator for check.

For example, if the user placed an order to sell his stock based on the current price. Let us assume a company's stock price was $75 at 10:20 am based on the information sent by the provider. The user sold his stocks based on that price. Because of the delay, the order was executed at the price of $65 since the correct price at 10:20 was $65. Therefore, she filed a claim and asked for a refund.

In this case, the provider should feed the system and the smart contract with a time-stamp along with the price. So the mediator will check and compare it with the current information. Eliminating third parties and exchanging transactions can be achieved using technologies, such as blockchain and smart contracts, which are described in the proposed system in Sect. 6.

6 Proposed Framework for a Decentralized Data Marketplace

With the use of decentralized data marketplaces, users can communicate with one another directly. Without the assistance of a middleman, payments and data are transferred and received simultaneously between the buyer and the seller. Figure 1 demonstrates the data marketplace's architecture. In this section, we are going to illustrate the proposed design.

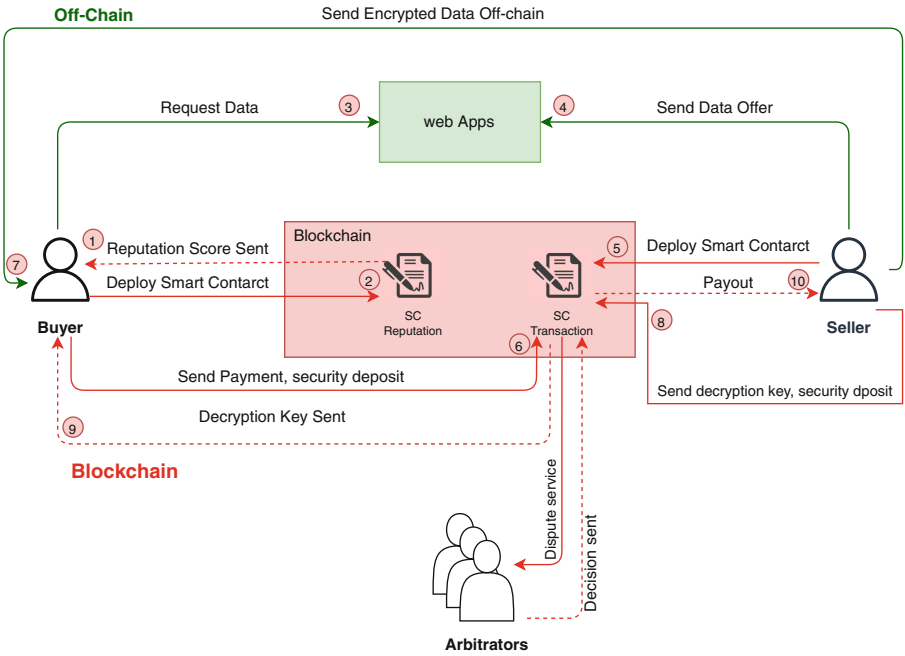

Fig. 1. A decentralized data marketplace framework

6.1 System Architecture

In this part, the component of the proposed decentralized data marketplace is presented.

Users: the marketplace has three main players listed below:

1. **Buyers** are a person or an entity who is willing to buy data from a seller.
2. **Sellers** are a person or an entity who owns data and willing to sell it to a buyer.
3. **Arbitrators** are an independent person who has been officially appointed to resolve a dispute.

Buyers could report to the authorities about any product in a centralized marketplace depending on the terms and conditions. As a result, they were able to return the things they had bought and get a refund. However, since there is no middleman between the buyer and the seller in a decentralized marketplace, this scenario is not applicable. Furthermore, it is impossible to return the data due to its nature. The buyer has the right to a refund in this situation if the requested data does not satisfy the contractual obligations of the buyer, but who would make that decision? Engaging arbitrators who can settle business conflicts in return for remuneration is one way to address this problem. Arbitrators' main

duty is to receive buyer disputes and settle them by looking into the contract. The smart contract is then informed of the outcome. Choosing arbitrators in a decentralized market is difficult for a number of reasons, including:

1. The selection of arbitrators should be agreed upon by both parties.
2. The arbitrators must be qualified and well-known.

Authors in [19] provide the Notaries approach as one way to choose arbitrators. They provide a system that enables the customer to pick a member from a list and extend an invitation to join. She then gives the vendors the list, who must select one before returning it to the customer after they have agreed. The notaries are to be paid under this agreement even if there is no dispute.

As an alternative, we choose arbitrators based on how well-known they are. Only when a dispute occurs using this method is the need for arbitrators present. In other words, the transaction is final if the buyer doesn't submit a claim. Thus, the arbitrators will only be compensated if a conflict arises. Both participants can select one arbitrator, and the smart contract will select one at random as a casting vote.

One problem that can arise is that if the cost of the data is low, consumers might decide not to file a complaint due to the high expense of doing so. We, therefore, have two categories of arbitrators to address this issue. Arbitrators who have more tokens than the minimum required by the data marketplace fall under the first type. These arbitrators will be given the security deposit that one of the parties would lose it. The second group of arbitrators has tokens below the threshold. These arbitrators will be encouraged to sign up as arbitrators so they may accumulate more tokens and reach the necessary level to start receiving payments.

In term of data price, the system selects arbitrators from the marketplace, assuming they are expert in the area.

- High price data– The system selects users randomly if their tokens are more than the $threshold_2$
- Low Price Data– The system selects users randomly no matter their tokens.

Web-Based Portal: the marketplace's front end is the web-based interface. A web or mobile application should be used by both sellers and buyers to submit their offers and requests. They may search, sell, and buy their data through the portal as well. Although the front-end is an off-chain marketplace component, it operates in conjunction with the blockchain components to carry out buyer and seller transactions.

Contract: is an agreements between two users (entities) that, when signing, creates commitments on each party. In the data marketplace, both users sign the contract using their private keys and hash it so they can check it. The contract must then be converted into a smart contract as the next step. Figure 2 shows the process of the contract before converted into a smart contract.

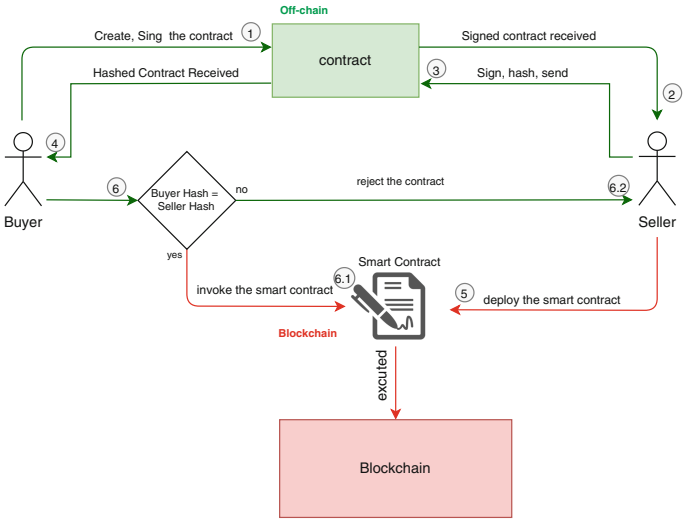

Fig. 2. The process of converting a contract to a smart contract

Smart Contract: because they are composed of rules defined in code, smart contracts are similar to conventional contracts in that regard. In other words, smart contracts are computer programs that run and exist at specified addresses on the Ethereum blockchain. So that they can transfer transactions throughout the network and maintain track of their money. But rather than being controlled by a person, they are placed on a network and executed as a program [5].

A smart contract serves as a trusted third party in the suggested system. Both buyers and sellers may rely on one another to execute transactions quickly and without additional costs. While exchanging data with the participants, it has the capacity to hold paid money. Additionally, the smart contract will use arbitrators to settle any disputes that could arise. Based on the arbitrator's decision, the smart contract completes the transaction. Additionally, the use of smart contracts aims to improve participant trust. Therefore, it's crucial to impose additional restrictions in order to stop people from misbehaving. To complete any transaction in this system, a security deposit is needed from both participants, which is a small payment known as a security deposit, made based on the users' agreement.

The purpose of the smart contract, as previously indicated, is to ensure that the buyer and seller are qualified to complete transactions. The smart contract guarantees that the buyer and seller fulfill the necessary conditions before the transaction can begin. In particular, the smart contract initiates the transaction and delivers the buyer the decryption key if the two hash values are equal. The smart contract starts a countdown timer *timer* after sending the URL to determine how much longer the transaction will take to finish in 24 h. Up until the timer's expiration, the dispute option is available. The buyer can begin and

decrypt it after receiving the URL of the encrypted data and the decryption key. The smart contract begins choosing the arbitrators if a dispute arises during the first 24 h of the transaction. The arbitrators then decide whether to accept or reject the dispute. Based on the transaction status, the transaction is considered successful if there is no dispute. Thus, the number of successful and total transactions is increased for both users.

Reputation System: authors in [20] define reputation as a measure of a participant's honesty based on past behavior. The reputation system in the data marketplace helps to build trust among participants by tracking their behavior and forcing them to behave honestly. Sellers with higher reputation scores may be able to sell their data at a higher price. Lower reputation score sellers, on the other hand, may find it difficult to compete with higher reputation score sellers. As a result, lowering the data price could be a solution to attracting buyers' attention and assisting them in rebuilding their reputation. In short, honest participants have market power based on their reputation score, whereas dishonest participants do not.

Reputation systems, on the other hand, have numerous flaws. One of the issues is that rating scores are exposed to other users, which discourages users from submitting negative ratings due to revenge [2]. Furthermore, users may choose not to provide feedback at all, or dishonest users may provide incorrect ratings.

As a result, a rating system based on user performance is one approach that can be derived from the proposed framework. In the proposed system, we can use transaction status to compute users' reputation scores based on their performance. The system provides an objective view of user performance in terms of the number of successful transactions and total transactions. To acquire a user's performance, reputation engines based on binary events or methods for rating aggregation based on simple mean and normal distribution can be implemented using a smart contract.

6.2 System Workflow

In this section, the framework workflow is presented. Figure 3 shows a successful transaction scenario where there is no dispute, ends with both the buyer and the seller total transaction increased by one transaction. Also, the total successful transaction increased by one transaction.

Figure 4 shows an unsuccessful transaction scenario where the buyer rises a dispute and it is accepted, ends with the seller's total and unsuccessful transaction increased by one.

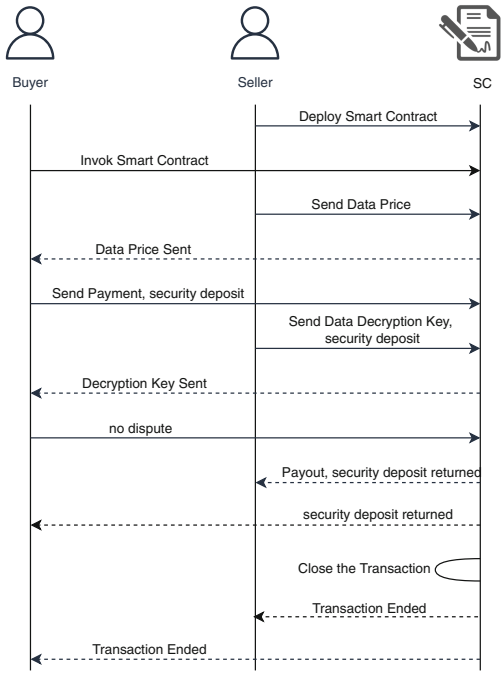

Fig. 3. A scenario for a successful data transaction in decentralized marketplace

In the following, a sequence of events and interactions flow for the proposed system:

1. The Buyer B request data $D_{is} = [DR, PK_B]$ and send it to Seller S to obtain
   (a) Data Requested DR
   (b) Buyer PK_s
2. The Seller S accept the request and send his data offer DO = P, MD, DH, EDU, TC, OH, D_{is} that include:
   (a) Data price P
   (b) Meta data MD
   (c) Data hash DH
   (d) URL Encrypted Data EDU
   (e) Terms and Conditions TC
   (f) Offer Hash OH
   (g) Data D_{is}
3. The buyer B accept the offer DO and send the following to the smart contract:
   (a) Data price P
   (b) Offer hash OH
   (c) Security Deposit SD_B

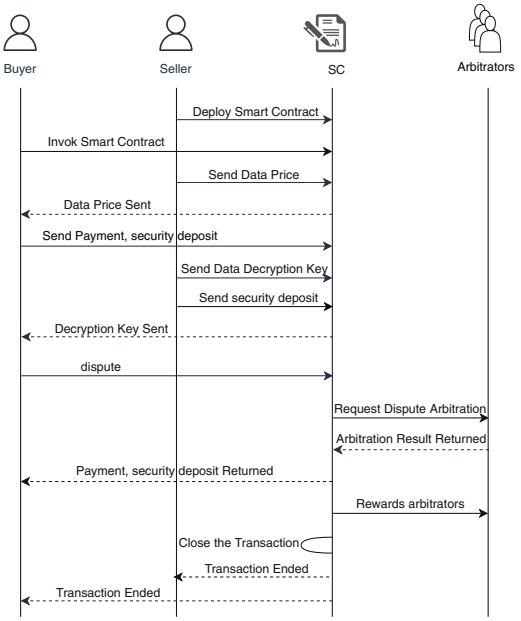

Fig. 4. A scenario for a unsuccessful data transaction in decentralized marketplace

4. The Seller S send the following to the Smart Contract:
   (a) Security Deposit SD_S
   (b) Data decrypted key DK
   (c) Offer hash OH
5. Smart Contract sends the following to the Buyer B
   (a) Data decrypted key DK
6. Smart Contract after Time T:
   (a) $CaseOne$ (NO dispute):
      i. Send the data Payment DP to the Seller S
      ii. Send the Security Deposit SD_S to the Seller S
      iii. Send the Security Deposit SD_B to the Buyer B
      iv. Submit Transaction Status for Seller S = Successful, Buyer B = Successful
   (b) $CaseTwo$ (dispute):
      i. Decision D = True, where buyer is right
         A. Send the data Payment to the Buyer B
         B. Send the Security Deposit SD_B to the Buyer B
         C. Send the Security Deposit SD_S to the Arbitrators
         D. Submit Transaction Status for Seller S = Unsuccessful, Buyer B = Successful
      ii. Decision D = False, where buyer is wrong
         A. Send the data Payment to the Seller S
         B. Send the Security Deposit SD_S to the Seller S

C. Send the Security Deposit SD_B to the Arbitrators
D. Submit Transaction Status for Seller S = Successful, Buyer B = Unsuccessful

7. Arbitrators receives the dispute from Buyer B and does the following:
   (a) Check on the dispute
   (b) Send the decision to the Smart Contract:
      i. Decision D = F, if the Buyer B is wrong
      ii. Decision D = T, if the Buyer B is correct

7 Towards a Global Data Marketplace

In e-commerce, such as data marketplaces, participants' locations could not be known, especially on a global scale. As a result, laws are unable to be enforced. Another legal problem for such global marketplaces appears to be determining which applicable legislation and jurisdiction to obey.

Several marketplaces set their rules and laws based on their nation of origin or as specified in their agreements, making them only subject to the laws of that country [14]. In Amazon, for example, users must agree to use the services based on the Federal Arbitration Act, applicable federal law, and the laws of the state of Washington. While on eBay, they must follow the law of the state of Utah or the American Arbitration Association ("AAA") in the case of arbitration.

In the proposed data marketplace, a seller and a buyer negotiate, then state their terms and conditions before submitting them to the smart contract. To solve this issue, the proposed system relies on the DRC, which works as an arbitration system. When a dispute is submitted, buyers can select one or two representatives from the marketplace as well as the sellers, or the system will select them instead. The representatives are assumed experts in the area. Therefor, in the event that the buyer files a claim, the arbitrators will study the claim and provide their decision based on the offer submitted to the smart contract.

8 Conclusion

In this work, a framework for a decentralized data marketplace based on blockchains and smart contracts is proposed. To prevent fraud, the system requires participants to provide a security deposit. In the event of a disagreement, the smart contract will refer the dispute to the DRC for resolution. The paper shows technologies such as blockchains and artificial intelligence being used to overcome challenges and concerns. Also, how the data market has evolved into a worldwide market is explored.

References

1. Alqwadri, A., Azzeh, M., Almasalha, F.: Application of machine learning for online reputation systems. Int. J. Autom. Comput. **18**(3), 492–502 (2021)
2. Bag, S., Azad, M.A., Hao, F.: A privacy-aware decentralized and personalized reputation system. Comput. Secur. **77**, 514–530 (2018)
3. Bar-Isaac, H.: Reputation and survival: learning in a dynamic signalling model. Rev. Econ. Stud. **70**(2), 231–251 (2003)
4. Bughin, J., Seong, J., James, M., Chui, M., Joshi, R.: Modeling the global economic impact of AI—mckinsey (2018). www.mckinsey.com/featured-insights/artificial-intelligence/notes-from-the-AI-frontier-modeling-the-impact-of-ai-on-the-world-economy. Accessed 06 Mar 2022
5. Introduction to smart contracts—ethereum.org (2021). https://ethereum.org/en/developers/docs/smart-contracts/. Accessed 14 Mar 2021
6. Hynes, N., Dao, D., Yan, D., Cheng, R., Song, D.: A demonstration of sterling: a privacy-preserving data marketplace. Proc. VLDB Endow. **11**(12), 2086–2089 (2018). https://doi.org/10.14778/3229863.3236266
7. Jha, V., Ramu, S., Shenoy, P.D., Venugopal, K.: Reputation systems: evaluating reputation among all good sellers. Data-Enabled Discov. Appl. **1**(1), 1–13 (2017)
8. Kim, M.S., Ahn, J.H.: Management of trust in the e-marketplace: the role of the buyer's experience in building trust. J. Inf. Technol. **22**(2), 119–132 (2007)
9. Miao, H., Li, A., Davis, L.S., Deshpande, A.: Towards unified data and lifecycle management for deep learning. In: 2017 IEEE 33rd International Conference on Data Engineering (ICDE), pp. 571–582. IEEE (2017)
10. Minhas, N.N.: Why iota is better and getting better than blockchains?—by noman nasir minhas—nerd for tech—medium. https://medium.com/nerd-for-tech/why-iota-is-better-and-getting-better-than-blockchains-b6eb62bf8b87. Accessed 17 July 2022
11. Pfeiffer, T., Nowak, M.A.: Digital cows grazing on digital grounds. Curr. Biol. **16**(22), R946–R949 (2006)
12. Prova, A.A., Akter, T., Islam, M.R., Uddin, M.R., Hossain, T., Hannan, M., Hossain, M.S.: Analysis of online marketplace data on social networks using LSTM. In: 2019 5th International Conference on Advances in Electrical Engineering (ICAEE), pp. 381–385. IEEE (2019)
13. Disclaimer (2022). https://www.reuters.com/info-pages/disclaimer/
14. Rosenthal, L., Mithal, M., Donohue, M., Sheer, A.: Consumer protection global electronic marketplace looking (2000). https://www.ftc.gov/sites/default/files/documents/reports/consumer-protection-global-electronic-marketplace-looking-ahead/electronicmkpl.pdf
15. Sharma, P., Lawrenz, S., Rausch, A.: Towards trustworthy and independent data marketplaces. In: Proceedings of the 2020 The 2nd International Conference on Blockchain Technology, pp. 39–45 (2020)
16. Soska, K., Kwon, A., Christin, N., Devadas, S.: Beaver: a decentralized anonymous marketplace with secure reputation. Cryptology ePrint Archive (2016)
17. Souza, R., et al.: Provenance data in the machine learning lifecycle in computational science and engineering. In: 2019 IEEE/ACM Workflows in Support of Large-Scale Science (WORKS), pp. 1–10. IEEE (2019)
18. Trabucchi, D., Buganza, T., Pellizzoni, E.: Give away your digital services: leveraging big data to capture value new models that capture the value embedded in the data generated by digital services may make it viable for companies to offer those services for free. Res. Technol. Manag. **60**(2), 43–52 (2017)

19. Travizano, M., Sarraute, C., Ajzenman, G., Minnoni, M.: Wibson: a decentralized data marketplace. arXiv preprint arXiv:1812.09966 (2018)
20. Witkowski, M., Artikis, A., Pitt, J.: Experiments in building experiential trust in a society of objective-trust based agents. In: Falcone, R., Singh, M., Tan, Y.-H. (eds.) Trust in Cyber-societies. LNCS (LNAI), vol. 2246, pp. 111–132. Springer, Heidelberg (2001). https://doi.org/10.1007/3-540-45547-7_7
21. Yoo, H., Ko, N.: Blockchain based data marketplace system. In: 2020 International Conference on Information and Communication Technology Convergence (ICTC), pp. 1255–1257. IEEE (2020)