# Towards a blockchain powered IoT data marketplace

Pooja Gupta*, Volkan Dedeoglu†, Salil S. Kanhere*, and Raja Jurdak‡
*UNSW, Sydney †CSIRO Data61, Brisbane ‡QUT, Brisbane
{pooja.gupta, salil.kanhere}@unsw.edu.au, volkan.dedeoglu@data61.csiro.au, r.jurdak@qut.edu.au

***Abstract*—The unprecedented rate of IoT adoption presents an opportunity for device owners to trade their IoT data with interested buyers. A blockchain-enabled data marketplace can democratize the trading of private IoT data by empowering data owners to choose what they want to share and with whom. However, some properties of IoT make it difficult to trade the generated data in conventional centralized markets. This research focuses on developing a marketplace framework to address design challenges imposed by IoT characteristics, such as limited resource and computational capabilities, mobility, data privacy and reselling issues. We propose a three-tiered framework to effectively tackle these challenges from elemental, functional and managerial aspects.**

***Index Terms*—Data marketplace, Internet of things, Blockchain**

## I. Introduction

The proliferation of IoT devices is generating an unprecedented volume of data, which is expected to facilitate a data economy of the market value of 3.9 trillion USD by 2030 [1]. New applications are emerging in every industry that promise to reduce costs, improve efficiency and productivity, and provide better service quality. However, most of the existing applications are designed to address specific problems, hence the generated data stay in silos. Moreover, instead of deploying new sensors and devices, government and organizations envision to employ already available infrastructure with normal citizens. Individuals can share data streams generated from their personal devices, smart homes, etc. Nonetheless, IoT data is underutilized as a large percentage of users are not willing to share their personal data due to the security and privacy concerns, and lack of monetizing mechanisms and incentives. To unleash the true potential of IoT data, an open data paradigm commonly known as data marketplace, is gaining popularity. A data marketplace enables data sellers to publish and sell their data and data buyers to purchase data based on their requirements.

To date, data marketplaces for trading IoT or non-IoT data have typically been developed using a centralized architecture. Given a large number of connected IoT devices and their generated data, having a centralized platform to govern and manage different aspects of data trading can be a bottleneck. Furthermore, centralized governance would be a costly option, in addition to other issues such as a single point of failure, mistrust, compromise and hacking. Therefore, there is a need to facilitate the trading of IoT data in a manner that is highly secure, scalable, trusted and decentralized. With its salient features such as immutability, decentralization, and enhanced security, blockchain is gaining significant traction. Blockchain smart contracts have the ability to disrupt the data marketplace and make trading more democratic, secure, autonomous, transparent, and efficient. Nevertheless, the adoption of blockchain to create an open IoT data marketplace is not straight-forward and requires design consideration from different perspectives.

In this research, we aim to design IoT data marketplaces of the future by considering design challenges from elemental, functional and managerial aspects. IoT presents a unique set of challenges at the device level, an elementary unit in the marketplace. IoT devices are generally resource-constrained in nature and that poses a challenge while serving multiple buyers' demands simultaneously without degrading user's experience. User privacy is another critical consideration since IoT data often contains sensitive information about their owners. IoT data is an intangible asset and can be resold independently without the owner's consent leading to monetary loss. Finally, it is advantageous to trade IoT data in near real-time. The second aspect of the research is to create a fully-functional and effective marketplace decentralized application (dapp) system, with no trusted third party involved in operational tasks. Key operational components of a marketplace include subscription management and execution, real-time support, data pricing, payment and settlement, and reputation mechanisms for entities to rate each other. Furthermore, the managerial perspective presents a unique set of challenges which include (i) managing product/query listing and identifying matches from billions of devices based on user requirements, (ii) mitigating fragmentation issue due to IoT device mobility, and (iii) imposing region-specific privacy regulations (e.g., GDPR or California Consumer Privacy Act).

## II. Related work

Various approaches for blockchain-enabled marketplace frameworks have been proposed. In [2], blockchain is used as a trade transaction management system that improves transparency and traceability. Making use of the immutability and distributed nature in [3], blockchain is used by sellers as a product catalogue that buyers can browse to find the proper data type based on their requirement. In other approach [4], blockchain smart contract is used to implement access control mechanism enabling the data owners to manage who can access their data and for how long.

The existing approaches fail to capture the IoT characteristics and design challenges as stated earlier. Before realizing the promise of blockchain for IoT data marketplace, scale issue around offers/queries listing and matching must be addressed. In particular, if we record the global product information on

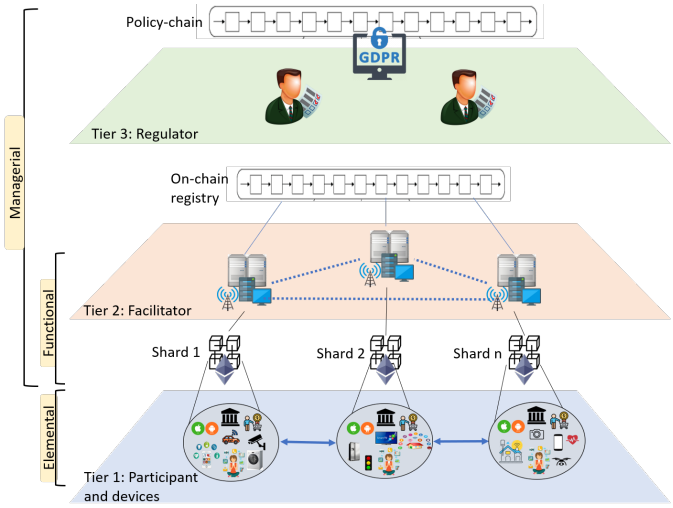


Fig. 1. 3-tier System architecture

the edge-devices, it can lead to an oversized chain and cause storage and scalability problems. Therefore, it becomes crucial to take these properties into account and develop a holistic architecture which can address complex IoT data marketplace requirements.

## III. Proposed Solution

In [5], we proposed a three-tiered system architecture as depicted in Figure 1. Tier 1 consists of all the participants, i.e. data sellers, their devices and data buyers. Tier 2 consists of geographically distributed facilitators that supervise specific service areas in return of incentives. Facilitators are realized using fog nodes to ease the burden on resource-constrained IoT devices. Facilitators use decentralized database systems such as BigChainDB to maintain the device and data information in a transparent and efficient manner. This enables a decentralized scheme for handling product information and segregates transaction requests dealing with references to the product/query data from the trading transaction. The trading related operational transactions are handled by marketplace decentralized user application, martchain. The third tier consists of regulators and auditors (e.g., data controllers as in GDPR) that make sure that the trading complies with regional regulations and policies. And since we are using blockchain technology for decentralization, it becomes important to ensure that privacy-sensitive data is not logged on to an immutable ledger. Supervisory authorities form a consortium blockchain network, policy-chain, and encode regulation policies in smart contracts. The purpose of regulators is to certify that underlying facilitators are continuously adhering to the regional privacy regulations. This tier ensures that data usage consent is met, only authorized buyers can process personal data, and audit data activities related to cross-regional trades. The contribution of this work is to develop a framework to tackle the design challenges as discussed previously using the following tiered approach.

### A. Elemental-level

IoT devices have restricted resource capabilities that limit the amount of data supply to serve multiple buyers simultaneously. A data supply is generally quantified in terms of sampling rate, duration and quality. These parameters directly impact the resource consumption of the IoT device. Furthermore, the motive of a seller is to maximize his revenue while the buyer's goal is to get quality data for the desired rate and duration. Incentive mechanisms based on pricing strategies can motivate users to trade their data. However, to guarantee the long-term engagement of participants and wider adoption of the marketplace, we focus on maximizing the participant satisfaction to create a sustainable marketplace design for trading real-time IoT data. To address this problem, we propose a two-step demand selection mechanism in [6]. The first step of matching happens at facilitator-end that matches a buyer's contextual requirement and seller's meta-data. While the second step of demand selection is performed on the seller's edge device based on the buyer's quantitative requirement as defined above and sellers' device's resource availability. This is achieved by using a multi-objective optimization framework that optimally selects demands to maximize both seller's revenue and the number of demands he can serve under various constraints of resources, allocation and quality.

### B. Functional-aspect

Martchain, a functional and effective marketplace, supports all the buying/selling activities as stated earlier. Additionally, it underpins mechanisms for (i) recording provenance of data to verify its genuineness, (ii) compensating data owners for unauthorized reselling of their data by other sellers, and (iii) enabling interoperability across other marketplace frameworks. Due to self-verifying, self-enforcing and tamper-proof characteristics, smart contracts can automate the execution flow of the trading whenever the predefined conditions are met without any administrative overhead. To meet the goal of developing an open platform, we employed the Ethereum public network. However, to execute smart contracts on Ethereum, transactions are submitted that consumes gas and hence, incurs cost. Therefore, developing a smart contract to implement the above components requires storage and processing optimization. In our proposed solution, as depicted in Figure 2, operational functionalities such as agreement management, data pricing and rating model are developed leveraging smart contracts. The agreement management controls all the aspect of trading starting from agreement creation, initiation, execution, payment and settlement. While existing approaches employ blockchain to record subscription or handle access control, we employ blockchain smart contract to automate the process flow of agreements. A new smart contract is deployed for agreement between each buyer/seller pair as a means to manage interdependent tasks. These dynamic contracts are customizable based on user-specification, maintains agreement integrity, enables autonomous data trading and ensures non-repudiation. The control flow of these smart contracts needs to be designed carefully to ensure the correct workflow of

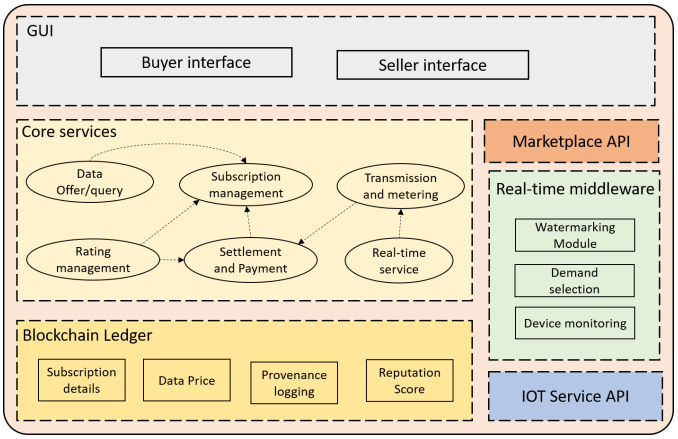


Fig. 2. Real-time IoT data Marketplace architecture

each stage of trading. The dynamic contract creation can result in an increased number of interconnected contracts over time. Therefore, managing the interactions of these contracts becomes challenging. Additionally, our proposed framework uses a watermarking technique to determine the reselling of data in [7]. We use a trade tracker contract that maintains the auditable trade trail and uses a fair payment scheme for sharing revenue among all the re-sellers. To enable interoperability and detection of unauthorized reselling across marketplace systems, we employ a digital notary that also provides proof of authenticity to buyers for any traded data in the system.

### C. Managerial-aspect

Our three-tiered architecture relies on multiple trust-less facilitators connected in a peer-to-peer fashion. The role of a facilitator is multifaceted. It manages product offers/requests, performs the discovery and match based on the active offers and recent queries, and resolves disputes caused by malicious and dishonest activities. To address scalability, we utilized cluster-sharding approach. Based on geographical locations, facilitator and participants in a service zone, are part of a sharded blockchain network. Meta-data of globally available sellers' offers are recorded in a separate ledger using on-chain registry maintained by facilitators. All the facilitators have a global view of the offers in the marketplace. When both the seller and the buyer are located in the same zone, the contracts are deployed in the same shard. When actors are residing in two different zones, cross-zonal contracts are deployed in the shards of both zones. However, this model suffers by a set of challenges including fragmentation, security and efficiency due to the mobility of IoT devices. As discussed earlier, the native shard, where the device is registered, is responsible for handling all the agreement related transactions. However, a device can visit other zones which can disrupt the on-going execution of active agreements and leads to fragmentation issue across various zones. To address this problem, we propose a contract handling mechanism for interzonal device movement. Based on the duration of the visit, the device registers itself as a temporary or permanent participant and contract-handover mechanism takes place. During permanent movement to a new zone, the device registers to the local facilitator and synchronizes with the local shard data. The seller can choose to either continue the active contracts by duplicating the agreements on the current shard or dissolve the agreement by paying the penalty. While for temporary visits, a short-term smart contract is deployed in the destination zone, which serves as a proxy contract for the on-going active contracts in other zones. Both the involved facilitators are incentivized for their contribution. Moreover, since the traded asset is privacy-sensitive data, regulators are responsible to audit and manage the trust-less facilitators to comply with the regional privacy regulations. Our design uses a trade-reconstruction mechanism in which regulators request facilitators to collect evidence by reconstructing random trades provided by regulators. A policy contract, encoded with native regulation policies, uses these pieces of evidence to validate if the trades are in compliance with encoded policies. Non-compliant facilitators are punished and penalized by the regulators.

## IV. Evaluation criteria

We will implement a proof-of-concept of our proposed architecture based on the Hyperledger Besu. Hyperledger Besu is an Ethereum based client supporting both public and permissioned network. We measure the following metrics for performance evaluation: (a) end-to-end delay, (b) monetary cost that the end-user has to pay as a transaction fee for trading, (c) throughput representing the number of trading processes that can be completed in each second, (d) blockchain size representing the memory footprint of the blockchain for storing all trading related transactions in each level.

## V. Conclusion

In this abstract, we highlighted design challenges to be considered in developing an IoT data marketplace framework. We proposed a 3-tier marketplace system addressing issues from different perspectives. Future work will focus on improving the reliability of the traded data by providing confidence estimates and employing an efficient authentication system for identifying and removing fake and malicious sellers from the market.

## References

[1] Gartner. 2020. Newsroom. [online] Available at: https://www.gartner.com [Accessed 4 November 2020].
[2] S. Bajoudah, C. Dong and P. Missier, "Toward a Decentralized, Trust-Less Marketplace for Brokered IoT Data Trading Using Blockchain," 2019 IEEE Blockchain, USA, 2019, pp. 339-346
[3] G. S. Ramachandran, R. Radhakrishnan and B. Krishnamachari, "Towards a Decentralized Data Marketplace for Smart Cities," 2018 IEEE ISC2, USA, 2018, pp. 1-8
[4] A. Suliman, Z. Husain, M. Abououf, M. Alblooshi and K. Salah, "Monetization of IoT data using smart contracts," in IET Networks, vol. 8, no. 1, pp. 32-37, 1 2019
[5] P. Gupta, S. S. Kanhere and R. Jurdak, A Decentralized IoT Data Marketplace, in Proceedings of 3rd SDLT, Gold Coast, November 2018.
[6] P. Gupta, V. Dedeoglu, K. Najeebullah, S. S. Kanhere and R. Jurdak, "Energy-aware Demand Selection and Allocation for Real-time IoT Data Trading," 2020 IEEE SMARTCOMP, Italy, 2020, pp. 138-147.
[7] P. Gupta, V. Dedeoglu, S. Kanhere, R. Jurdak, "Data Reselling in IoT Data Marketplace," In proceedings of the 7th ACM Celebration of Women in Computing: womENcourage 2020 conference, Azerbaijan, September, 2020.