ISSN: 2178-7727

Blockchain-Enabled Scalable Data Marketplaces for the Internet of Things

Dhafer Sabah Yaseen[1], Dr. Hardik Joshi[2]

[1]Department of Computer Sciences, University of AL-Hamdaniya, Ninevah, Iraq.
Email: dhafer.sabah@uohamdaniya.edu.iq
[2]Department of Computer Science, Gujarat University, Ahmedabad, India.
Email: hardikjoshi@gujaratuniversity.ac.in

ABSTRACT

As the Internet of Things (IoT) expands, so does the volume of valuable data it generates. Yet most of this data remains siloed, underutilized, or monopolized by platform providers. A decentralized data sharing infrastructure is needed to ensure equitable access, trust, and economic incentive for data exchange. This paper introduces the Blockchain-based IoT Data Marketplace (BIDM), a scalable, decentralized platform that allows IoT data producers and consumers to engage in secure and trustworthy data transactions. Built on Ethereum with a Proof-of-Authority consensus mechanism and leveraging smart contracts, BIDM ensures data ownership, fair pricing, secure access control, and transparent transaction logging. It supports both individual data purchases and stream-based subscriptions, offering flexibility to diverse consumers. The system is rigorously evaluated through synthetic and real-world IoT workloads, including integration with the SmartSantander platform. Performance results demonstrate the BIDM's viability under high loads, with consistent execution times and minimal scalability impact. This work provides a significant step toward secure, interoperable, and efficient blockchain-enabled data marketplaces for the IoT ecosystem.

KEYWORDS: Internet of Things, Data Marketplace, Blockchain, Smart Contracts, Data subscription, Performance evaluation

INTRODUCTION

The data economy is rapidly becoming a central pillar of global digital transformation. In the European Union alone, it is projected to contribute €800 billion to the GDP by 2025. Yet, a major bottleneck persists: vast amounts of data, particularly from IoT devices, are either inaccessible, underutilized, or locked within proprietary platforms. Despite IoT's potential to drive innovation across smart cities, health, transportation, and manufacturing, its data remains siloed due to concerns over privacy, lack of interoperability, and absence of incentive structures.

Traditional IoT solutions are fragmented, vertically integrated, and reliant on centralized infrastructures. This leads to vendor lock-in, poor scalability, and limited data reuse across applications. Moreover, many IoT data owners are reluctant to share their data due to trust issues and lack of control over how their data is used or monetized.

To overcome these limitations, there is a need for a secure-by-design, decentralized data sharing infrastructure that shifts the current trust model away from central authorities toward liquid, distributed trust. Blockchain technology, with its inherent transparency, immutability, and support for smart contracts, offers a promising foundation.

This paper proposes the Blockchain-based IoT Data Marketplace (BIDM), a decentralized platform that enables IoT data producers to share and monetize their data while maintaining ownership, access control, and trust. BIDM addresses key technical and governance challenges in building such a platform and presents a performance evaluation that demonstrates its viability in real-world settings.

ISSN: 2178-7727

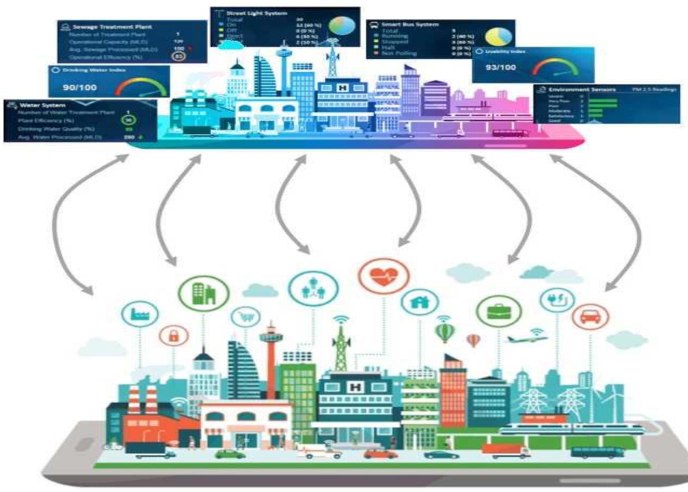

Fig. 1. Conceptual running case for a Smart City Digital Twin.

RESEARCH OBJECTIVES

To guide the development and evaluation of BIDM, the paper addresses the following research questions:

- **RQ1**: How can blockchain technologies support flexible and trustworthy exchange of IoT data?
    - What limitations exist in current data marketplaces?
    - What blockchain features can address these limitations?
    - What procedures are needed to operationalize a decentralized marketplace?
- **RQ2**: What is the performance overhead introduced by integrating blockchain into IoT data marketplaces?
    - ✓ How does blockchain affect execution time and system scalability?
    - ✓ How does system performance evolve under increasing load?

RELATED WORK

Centralized Marketplaces and Limitations

Early data marketplace models evolved from the Sensing-as-a-Service (SaaS) paradigm, where IoT data is aggregated in central cloud platforms. Such platforms dominate the marketplace but suffer from inherent drawbacks:

- Data producers have limited control over pricing and data exposure.
- Consumers pay platform fees, raising costs.
- Trust depends on a single platform authority.
- Platform operators can access all data, creating privacy risks.

Studies [9–13] have reviewed such architectures, but most fail to address decentralization, privacy, and fair governance.

Decentralized Marketplaces and Blockchain Proposals

Recent efforts have shifted toward decentralized models using blockchain for immutable logging, smart contracts, and trustless interactions. However, many of these efforts are conceptual or lack full implementations and empirical validation. Notable projects include:

- Data rating systems [22, 23]
- Privacy-focused exchanges [21, 25]
- Monetization and reputation mechanisms [26, 27]

Contribution of BIDM

BIDM stands out by offering:

- Full-stack implementation using Ethereum + IPFS
- Smart contracts for access control, pricing, and subscriptions
- Support for both individual measurements and stream-based subscriptions
- A complete performance evaluation using synthetic and real data (SmartSantander)

ISSN: 2178-7727

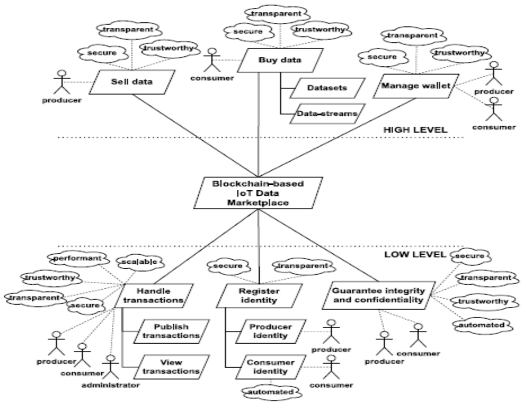

Fig. 2. AOM of the BIDM's requirements.

ARCHITECTURE AND SYSTEM DESIGN

The Blockchain-based IoT Data Marketplace (BIDM) is designed to facilitate trustworthy, scalable, and decentralized exchange of IoT data. The platform architecture integrates Ethereum-based smart contracts, decentralized off-chain storage, and lightweight APIs to enable seamless participation for both data producers and consumers.

System Requirements and Design Principles

BIDM's design addresses the following key requirements:

1. Decentralized Trust: Replace platform-centric authority with a trust model based on blockchain consensus.
2. Access Control: Enforce permissioned access to IoT measurements.
3. Verifiability: Provide immutable records of data ownership, pricing, and consumption.
4. Scalability: Handle high-frequency data operations without performance degradation.
5. Producer Autonomy: Allow producers to control pricing, visibility, and data availability.
6. Operational Flexibility: Support both one-time purchases and real-time stream subscriptions.

These goals are achieved by combining permissioned blockchain with modular smart contracts, off-chain storage via IPFS, and modular interfaces for IoT producers and data consumers.

Core Architecture Components

Blockchain Layer

BIDM operates on an Ethereum blockchain using the **Proof-of-Authority (PoA)** consensus algorithm (specifically, Clique). This choice reduces energy consumption, provides predictable block times, and avoids gas fees, essential for high-frequency IoT use cases.

Validators are approved nodes that maintain consensus. The blockchain logs:

- Measurement metadata
- Purchase transactions
- Subscription records
- User account registrations

Smart Contracts

There are three key smart contracts:

- Data Smart Contract (Data SC): Registers observations, links them to "Topics", and stores encrypted content identifiers (CIDs) pointing to off-chain IPFS data.
- Access Smart Contract (Access SC): Manages identity and device whitelisting. It ensures only verified producers can register data.
- Clearing House Smart Contract: Handles purchases and token transfers, records access rights, and enables subscription billing.

All contracts are deployed once and referenced by clients and backend services via Ethereum addresses.

ISSN: 2178-7727

Off-Chain Storage

IoT measurements are stored in IPFS (InterPlanetary File System). To ensure privacy and ownership:

- Each Topic has a symmetric encryption key.
- Data is encrypted before being stored in IPFS.
- Consumers receive the decryption key only upon successful purchase or a valid subscription.

This design minimizes blockchain bloat while preserving data immutability and decentralization.

Market Module

The Market Module has two subcomponents:

- **IoT Proxy**: Serves as an API gateway for producers. It handles authentication, encryption, and metadata extraction.
- **Market Core**: Executes administrative functions, listens to blockchain events (e.g., purchases), and coordinates data delivery.

Though central in operation, this module does not introduce centralization, as its actions are logged on-chain and verifiable.

Client Interfaces

BIDM exposes a RESTful web API for producers and consumers. This interface:

- Registers new accounts and devices
- Enables users to list available Topics and datasets
- Handles purchases and subscription flows

The user experience is designed to abstract blockchain complexity while preserving transparency and accountability.

DATA EXCHANGE PROCEDURES

The platform supports two main data acquisition modes: individual purchase and **subscription**. Both modes share a foundational pipeline involving smart contracts, metadata registration, token transfers, and IPFS-based retrieval.

Measurement Registration

The procedure begins when a data producer submits a new measurement via the IoT Proxy. Key steps:

1. Preprocessing: Metadata (timestamp, Topics, etc.) and credentials are extracted.
2. Encryption & Storage: Measurement is encrypted with Topic-based keys and uploaded to IPFS. The CID is returned.
3. On-chain Registration:
   - ➢ Data SC logs the measurement hash, Topics, and encrypted CID.
   - ➢ Clearing House SC logs its price.
4. Notification: If the measurement matches an active subscription, the consumer is notified.

This workflow supports stream tagging, enabling later subscription.

One-Time Purchase

A consumer initiates a purchase by selecting a measurement and invoking a transaction via the Clearing House SC.

Steps:

1. **Token Transfer**: The Consumer pays the BIDM administrator wallet.
2. **Blockchain Logging**: Purchase details (buyer/seller, measurement hash) are logged immutably.
3. **Data Delivery**:
   - ➢ Market Core retrieves and encrypts the CID with the buyer's public key.
   - ➢ The consumer retrieves and decrypts data from IPFS.

A final confirmation logs the transaction hash and access credentials.

ACTA SCIENTIAE
ISSN: 2178-7727

Topic-Based Subscription
Consumers may subscribe to Topics representing streams of data (e.g., "temperature in downtown sensors").
Steps:
1. Topic Browsing: Consumers browse available Topics via the API.
2. Subscription Request: A one-time token transfer is recorded on-chain.
3. Notification Pipeline: New matching measurements are automatically pushed to the subscriber's wallet, encrypted with the shared Topic key.

Subscriptions minimize on-chain transactions, making this model more efficient at scale.

Model Comparison

Feature	Individual Purchase	Subscription
Access Type	Historical/specific	Real-time stream
Token Cost	Pay per item	Flat per-topic
Scalability	Less scalable	Highly scalable
Use Case	Targeted data	Continuous monitoring
Performance	Three on-chain TXs per item	One-time TX only

BIDM allows both models to co-exist, offering flexibility based on consumer intent and application needs

PERFORMANCE EVALUATION
To validate the feasibility of BIDM under realistic workloads, a comprehensive performance evaluation was conducted using both synthetic data generators and a real-world IoT deployment (SmartSantander). The goal was to assess execution times, scalability, and system stability under high loads.

Experimental Setup
- **Blockchain Configuration**: Ethereum PoA test network (Clique) with 3–6 nodes.

✓ Roles: Validator, Market Node, Client Node

- **Measurement Sources**:

✓ Synthetic data generator with adjustable inter-arrival times and payload size
✓ Real-world data from SmartSantander's smart city sensors

- **Metrics Evaluated**:
  - ✓ Execution time for data registration, purchase, and subscription
  - ✓ Scalability under increasing throughput
  - ✓ Impact of validator count
  - ✓ Disk space consumption (footprint)

Registration Time Performance
Fixed vs Poisson Arrival
Measurements were generated at:
- Fixed intervals (2s, 1s, 0.5s, 0.2s)
- Poisson-modeled bursts (mean 0.5s)

Findings:
- **Average registration time**: ~130 ms
- **Performance unaffected** by data bursts or shorter intervals
- The system handled over 500,000 registrations without degradation

Effect of Data Size
- Measurement sizes tested: 55B, 229B, 997B
- No significant impact on registration time from payload size

ISSN: 2178-7727

Real-World Integration (SmartSantander)
- 10,000 real measurements registered
- Execution times comparable to synthetic tests
- Validates practical integration with live IoT systems

Impact of Validator Nodes
- Compared configurations with 1, 3, and 6 validators
- Performance varied slightly (±15%) but remained stable
- Scalability is preserved as network decentralization increases

Purchase Performance
- Average purchase time: ~100–120 ms
- Poisson bursts showed faster response due to batch mining behavior
- Measurement size did not affect performance
- Consistent results with real and synthetic data

Subscription Model Performance
- Initial subscription time: ~140 ms
- No further transactions needed per measurement
- Ideal for high-frequency data consumers (e.g., real-time analytics)
- Outperforms individual purchases in long-term cost and execution overhead

Storage Footprint and Stability
- Only metadata is stored on-chain; full data resides in IPFS
- The footprint remains small and predictable
- System stability is preserved with long-running sessions and heavy load

CONCLUSION AND FUTURE WORK

Summary of Contributions

This paper presents BIDM, a decentralized, scalable data marketplace designed to unlock the potential of IoT data sharing while preserving ownership, transparency, and control. Key features include:
- Permissioned blockchain (Ethereum PoA) for fast, low-cost consensus
- Smart contracts manage data access, payments, and subscriptions
- Encrypted off-chain storage using IPFS
- Support for both synchronous purchases and asynchronous subscriptions
- Real-world evaluation under synthetic and smart city workloads

The implementation and empirical validation show that BIDM achieves its goals:
- Minimal execution time per transaction (~130ms)
- No performance degradation under high load
- Modular and extensible architecture

Implications and Use Cases

BIDM is applicable in any domain where real-time or historical IoT data is of value but trust, privacy, and incentives are lacking. Potential use cases include:
- Smart cities
- Environmental monitoring
- Smart transportation
- Industrial IoT

Its flexible architecture enables interoperability between producers and consumers, allowing SMEs, researchers, and enterprises to participate without the need for intermediaries.

ISSN: 2178-7727

Future Work

The authors identify several directions to extend this work:

- Reputation mechanisms for rating producers and consumers
- Semantic Topic standardization to avoid fragmentation
- Tokenomics model and integration with external cryptocurrency systems
- Cross-chain data marketplaces enabling federated interoperability
- Dynamic access control policies based on context or user profile

Further studies may also include economic simulations to optimize data pricing and subscription strategies.

CONCLUSION

This paper has presented the Blockchain-based IoT Data Marketplace (BIDM) as a comprehensive solution for enabling secure, decentralized, and scalable exchange of IoT-generated data. By leveraging blockchain technology specifically Ethereum with a Proof-of-Authority consensus mechanism alongside smart contracts and off-chain storage, BIDM facilitates a transparent and trusted data marketplace where producers maintain control over their data assets and consumers gain verified, timely access.

The dual-mode acquisition approach offering both individual measurement purchases and Topic-based subscriptions—caters to diverse application needs, ranging from targeted historical analysis to continuous real-time monitoring. Through its design, BIDM addresses key challenges in data governance, privacy, and market fairness, which are critical to overcoming the current limitations of centralized IoT data platforms.

The implementation and performance evaluation conducted under both synthetic and real-world IoT workloads (including the SmartSantander smart city infrastructure) demonstrate that BIDM meets not only functional requirements but also crucial non-functional ones such as execution time, resource efficiency, and scalability. With average transaction times in the 100–130 ms range and consistent behavior under increasing system load, BIDM proves to be a robust candidate for deployment in demanding IoT environments.

Beyond the technical validation, BIDM offers a model for open, trust-based data economies—particularly relevant for smart cities, industrial IoT, and research ecosystems—where data can be shared securely and monetized fairly without reliance on centralized authorities. The platform supports interoperability, extensibility, and dynamic governance, laying a solid foundation for future integration with cross-domain data spaces, reputation systems, and token-based incentive mechanisms.

In conclusion, this work marks a significant advance in the design and practical realization of decentralized IoT data marketplaces. BIDM not only validates the feasibility of blockchain-based solutions for complex data ecosystems but also provides a blueprint for scalable, sustainable, and secure data economy infrastructures.

References

1. K.J. Singh, D.S. Kapoor, Create your own internet of things: A survey of IoT platforms, IEEE Consum. Electron. Mag. 6 (2) (2017) 57-68, http://dx.doi.org/10.1109/MCE.2016.2640718.
2. M. Noura, M. Atiquzzaman, M. Gaedke, Interoperability in internet of things: Taxonomies and open challenges, Mob. Netw. Appl. 24 (3) (2019) 796-809, http://dx.doi.org/10.1007/s11036-018-1089-9.
3. P. Sotres, J. Lanza, L. Sanchez, J.R. Santana, C. Lopez, L. Munoz, Breaking vendors and city locks through a semantic-enabled global interoperable internet-of-things system: A smart parking case, Sensors 19 (2) (2019) 229, http://dx.doi.org/10.3390/s19020229.
4. C. Harrold, Beyond the hype-smart today and tomorrow, in: Practical Smart Device Design and Construction, Springer, 2020, pp. 37-45, http://dx.doi.org/10.1007/978-1-4842-5614-5.
5. L. Sanchez, J. Lanza, L. Munoz, From the internet of things to the social innovation and the economy of data, Wirel. Pers. Commun. 113 (3) (2020) 1407-1421, http://dx.doi.org/10.1007/s11277-020-07321-2.
6. X. Zhao, K.K. Sajan, G.S. Ramachandran, B. Krishnamachari, Demo abstract: the intelligent IoT

ISSN: 2178-7727

integrator data marketplace-version 1, in: 2020 IEEE/ACM Fifth International Conference on Internet-of-Things Design and Implementation (IoTDI), IEEE, 2020, pp. 270-271, http://dx.doi.org /10.1109/IoTDI49375.2020.00042.
7. Agarwal, M. Dahleh, T. Sarkar, A marketplace for data: An algorithmic solution, in: Proceedings of the 2019 ACM Conference on Economics and Computation, 2019, pp. 701-726, http://dx.doi.org/10.1145/3328526.3329589.
8. P. Gupta, S. Kanhere, R. Jurdak, A decentralized IoT data marketplace, 2019, http://dx.doi.org/10.48550/arXiv.1906.01799, arXiv preprint arXiv:1906.01799.
9. H. Yoo, N. Ko, Blockchain based data marketplace system, in: 2020 International Conference on Information and Communication Technology Convergence, ICTC, IEEE, 2020, pp. 1255-1257, http://dx.doi.org/10.1109/ICTC49870.2020.9289087.
10. P. Banerjee, S. Ruj, Blockchain enabled data marketplace-design and challenges, 2018, http://dx.doi.org/10.48550/arXiv.1811.11462, arXiv preprint arXiv:1811.11462.
11. S. Lawrenz, P. Sharma, A. Rausch, Blockchain technology as an approach for data marketplaces, in: Proceedings of the 2019 International Conference on Blockchain Technology, 2019, pp. 55-59, http://dx.doi.org/10.1145/3320154.3320165.
12. P. Gupta, V. Dedeoglu, S.S. Kanhere, R. Jurdak, Towards a blockchain powered IoT data marketplace, in: 2021 International Conference on Communication Systems & NETworkS, COMSNETS, IEEE, 2021, pp. 366-368, http://dx.doi.org/10.1109/COMSNETS51098.2021.9352865.
13. G.S. Ramachandran, R. Radhakrishnan, B. Krishnamachari, Towards a decentralized data marketplace for smart cities, in: 2018 IEEE International Smart Cities Conference (ISC2), IEEE, 2018, pp. 1-8, http://dx.doi.org/10.1109/ISC2.2018.8656952.
14. M. Ha, S. Kwon, Y.J. Lee, Y. Shim, J. Kim, Where WTS meets WTB: A blockchain-based marketplace for digital me to trade users' private data, Pervasive Mob. Comput. 59 (2019) 101078, http://dx.doi.org/10.1016/j.pmcj.2019.101078.
15. R. Xu, G.S. Ramachandran, Y. Chen, B. Krishnamachari, Blendsm-ddm: Blockchain-enabled secure microservices for decentralized data marketplaces, in: 2019 IEEE International Smart Cities Conference (ISC2), IEEE, 2019, pp. 14-17, http://dx.doi.org/10.1109/ISC246665.2019.9071766.
16. N. Hynes, D. Dao, D. Yan, R. Cheng, D. Song, A demonstration of sterling: a privacy-preserving data marketplace, Proc. VLDB Endow. 11 (12) (2018) 2086-2089, http://dx.doi.org/10.14778/3229863.3236266.
17. Alsharif, M. Nabil, A blockchain-based medical data marketplace with trustless fair exchange and access control, in: GLOBECOM 2020-2020 IEEE Global Communications Conference, IEEE, 2020, pp. 1-6, http://dx.doi.org/10.1109 /GLOBECOM42002 .2020 .93 4 8192.
18. V. Koutsos, D. Papadopoulos, D. Chatzopoulos, S. Tarkoma, P. Hui, Agora: a privacy-aware data marketplace, IEEE Trans. Dependable Secure Comput. (2021) http://dx.doi.org/10.1109/TDSC.2021.3105099.
19. W. Badreddine, K. Zhang, C. Talhi, Monetization using blockchains for IoT data marketplace, in: 2020 IEEE International Conference on Blockchain and Cryptocurrency, ICBC, IEEE, 2020, pp. 1-9, http://dx.doi.org/10.1109/ICBC48266.2020.9169424.
20. S.A. Azcoitia, N. Laoutaris, A survey of data marketplaces and their business models, 2022, http://dx.doi.org/10.48550/arXiv.2201.04561, arXiv preprint arXiv:2201.04561.
21. F. Schomm, F. Stahl, G. Vossen, Marketplaces for data: an initial survey, ACM SIGMOD Record 42 (1) (2013) 15-26, http://dx.doi.org/10.1145/2481528.2481532.
22. M. Spiekermann, Data marketplaces: Trends and monetisation of data goods, Intereconomics 54 (4) (2019) 208-216, http://dx.doi.org/10.1007/s10272-019-0826-z.
23. L. Sterling, K. Taveter, The Art of Agent-Oriented Modeling, MIT Press, 2009.
24. Norta, A. Kormiltsyn, C. Udokwu, V. Dwivedi, S. Aroh, I. Nikolajev, A blockchain implementation for configurable multi-factor challenge-set self¬sovereign identity authentication, in: 2022 IEEE International Conference on Blockchain (Blockchain), 2022, pp. 455-461, http://dx.doi.org/10.1109/Blockchain55522.2022.00070.
25. F. Metzger, T. HoEfeld, A. Bauer, S. Kounev, P.E. Heegaard, Modeling of aggregated IoT traffic and its application to an IoT cloud, Proc. IEEE 107 (4) (2019) 679-694,

ISSN: 2178-7727

http://dx.doi.org/10.1109/JPROC.2019.2901578.
26. B. Otto, S. SteinbuE, A. Teuscher, S. Lohmann, Reference architecture model for the international data spaces, Tech. Rep., International Data SpacesAssociation, 2019.