Environment Systems and Decisions
https://doi.org/10.1007/s10669-023-09907-0

Decentralized brokered enabled ecosystem for data marketplace in smart cities towards a data sharing economy

Bokolo Anthony Jr[1,2]

Accepted: 16 March 2023

Abstract
Presently data are indispensably important as cities consider data as a commodity which can be traded to earn revenues. In urban environment, data generated from internet of things devices, smart meters, smart sensors, etc. can provide a new source of income for citizens and enterprises who are data owners. These data can be traded as digital assets. To support such trading digital data marketplaces have emerged. Data marketplaces promote a data sharing economy which is crucial for provision of available data useful for cities which aims to develop data driven services. But currently existing data marketplaces are mostly inadequate due to several issues such as security, efficiency, and adherence to privacy regulations. Likewise, there is no consolidated understanding of how to achieve trust and fairness among data owners and data sellers when trading data. Therefore, this study presents the design of an ecosystem which comprises of a distributed ledger technology data marketplace enabled by message queueing telemetry transport (MQTT) to facilitate trust and fairness among data owners and data sellers. The designed ecosystem for data marketplaces is powered by IOTA technology and MQTT broker to support the trading of sdata sources by automating trade agreements, negotiations and payment settlement between data producers/sellers and data consumers/buyers. Overall, findings from this article discuss the issues associated in developing a decentralized data marketplace for smart cities suggesting recommendations to enhance the deployment of decentralized and distributed data marketplaces.

Keywords Emerging technologies · Decentralized data marketplace · IOTA technology · MQTT broker · Data ecosystem · Data sharing economy · Smart cities · Sustainable society

1 Introduction

Data are a key component for success in many sectors hence it is being monetized (Alvsvåg et al. 2022). In smart cities data are generated from different sources such as vehicles, air quality monitors, smart trash cans, security, and surveillance, etc. which improves the lives of citizens. However, despite all the advantages derived from data, they are rarely traded or shared on a large scale (Hatamian 2021). Accordingly, there is need to design data marketplaces that can leverage available data from different sources to develop innovative applications (Hatamian 2021). A data marketplace is an online store or digital platform that facilitates data trading for buyers and sellers. Data marketplace aids data consumers which may be service providers or application developers to compensate and incentivize data owners for sharing data fostering a data sharing platform where data owners, sellers and buyers can sell, find, and buy data as easily as possible (Ramachandran et al. 2018). Centralized data marketplaces already exist but require a central authority or trusted third party to negotiate the data exchange. This introduces trust concerns, as the central authority generally has access to the data source traded and may illegally benefit from it. Such concerns may prevent data owners from being part of a data marketplace since their data may not be secured (Meijers et al. 2021; Barroso and Laborda et al. 2022).

Findings from the literature explored how a decentralized data marketplace could be developed using emerging technologies such as distributed ledger technology (DLT)

✉ Bokolo Anthony Jr
anthony.j.bokolo@hiof.no; anthony.bokolo@ife.no

[1] Department of Applied Data Science, Institute for Energy Technology, 1777 Halden, Norway

[2] Department of Computer Science and Communication, Østfold University College, BRA Veien 4, 1757 Halden, Norway

or blockchain (Ramachandran et al. 2018; Xu et al. 2019; Bajoudah and Missier et al. 2021; Hatamian 2021). DLTs such as blockchains mostly deploys a decentralized peer-to-peer (P2P) consensus network without the need for a centralized trusted third-party authority (Marikyan et al. 2022). The decentralization architecture.Employed by DLT helps to improve security, mitigates performance issues and eliminates any single point of failure which is inherent within central hub such as in data management systems. Also, DLT leverage consensus protocols to support traceability, auditability, and immutability for data provenience (Xu and Chen 2021). Therefore, IOTA tangle is proposed to be used in this study as an emerging technology. IOTA similar to other DLT such as Tendermint and HyperLedger Fabric does not require fees as compared to DLTs such as Ethereum which transaction fees can be expensive (the gas costs associated with Ethereum transactions) (Meijers et al. 2021). Such fees can be a barrier to smaller payments such as those frequently required for trading IoT data streams in smart cities (Meijers et al. 2021).

By using IOTA tangle data producers or data owners/data sellers and data buyers/data consumers can securely enters into agreements without any intervention of trusted third party achieving a secure trading mechanism that allows the safe trading of data sources. Additionally, in this study the message queuing telemetry transport (MQTT) protocol is employed. MQTT utilizes a brokerage concept to broadcast information as messages (Ramachandran et al. 2019). In smart city domain MQTT is a generally utilized to transfer Internet of Things (IoT) data in real time (Meijers et al. 2021), based on the publish-subscribe messaging method, in which a broker is employed to orchestrate data communication among data producers and consumers (Anthony Jnr et al. 2020). This study is an answer to the call from Ramachandran et al. (2018), where the authors advocated for research that provides an architectural perspective to explore how a data marketplace can be decentralized. To accomplish this goal this study aims to address these following research questions;

- To identify the components needed to support a data marketplace in smart cities and how these components could be deployed without a trusted third parties or centralized servers.
- To design an ecosystem for a decentralized broker-based data marketplace that could potentially instill trust and fairness among data owners and data sellers when trading data in smart cities.

Therefore, this article designs an ecosystem for data marketplaces to facilitates critical operational and business decisions for application developers and also aids in achieving trust and fairness among different stakeholders involved in digital data trading within smart cities as this is not well researched in the literature. The designed ecosystem is powered by IOTA technology as a DLT and MQTT to facilitate trust and fairness among data owners and data sellers. The remainder of this article is structured as Sect. 2 provides the literature review. Section 3 describes the method which presents the designed ecosystem for data marketplaces, data collection and analysis method. Section 4 presents the findings and Sect. 5 is the discussion and implications from this work and finally, conclusions, limitations, and future works are presented in Sect. 6.

2 Literature review

This section mainly provides theoretical understanding of data and centralized/decentralized data marketplaces in smart cities. Besides, prior studies that investigated issues related to data marketplaces in smart cities are discussed.

2.1 Usefulness of data in making cities smarter

The European Commission announced in 2012 that smart cities would become one of the facilitators of sustainable development. Due to increased environmental awareness in society todays, cities are adopting implementing sustainable initiatives (Delgado-de Miguel et al. 2019). Smart cities explore the economic, social, environmental, and technological aspects of the urban space as seen in Fig. 1.

Figure 1 depicts the aspect of sustainability in smart cities. A smart city improves the citizen`s quality of life by

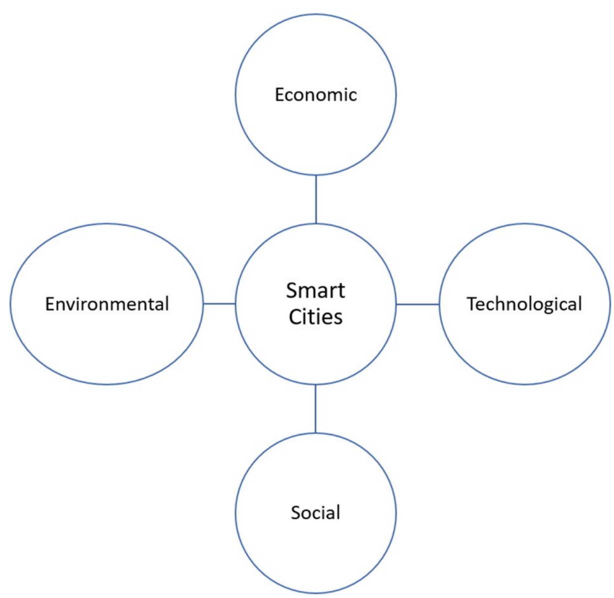


Fig. 1 Aspect of sustainability in smart cities adopted from Anthony Jnr (2021c)

addressing city problems (Jeong et al. 2020), as related to the aspects captured in Fig. 1 (Anthony Jnr 2021c). A smart city is simply a city that is uses novel information and communications technology (ICT) strategically and innovatively to achieve the aims of the municipality (Jnr and Petersen 2022). Smart cities aim to better control and manage urban systems by using available data real time, historic, online, etc. to optimize decision making for short-, medium-, and long-term sustainability goals (Cosgrave et al. 2013; Jnr et al. 2020). Smart city comprises of system of systems, which includes an ecosystem of services (Legenvre et al. 2022), people, resources, products, systems, enterprises, and society that are working together innovatively to foster sustainable development of the city (Kieti et al. 2021). Smart cities provide places that encourage creativity, where citizens are creators of ideas, solutions, and services rather than passive beneficiaries of services (Cosgrave et al. 2013). Over the years the environmental and social aspects have been researched, but the connection between economic and technology aspects has seen a limited number of studies (Vaninsky 2021). Research on the integration of the economic and technology aspects to improve the quality of life of citizens by meeting the social needs of present and future generations towards urban development is required (Delgado-de Miguel et al. 2019).

Data can help in digitalization of cities into smart cities (Jeong et al. 2020; Anthony Jnr 2021b). The proliferation of data aids the actualization of smart cities to become feasible by facilitating data driven services which have significantly improved the lives of citizens by improving safety, health, and convenience (Xu et al. 2019). Availability of relevant data supports municipality administration create effective strategies in making better decisions regarding urban development (Lawrenz et al. 2019). In smart cities there are individual data owners, each owning a couple of devices (Mišura and Žagar 2016). These smart devices and meters produce massive amounts of data in urban environment that provide valuable information which can be utilized to improve citizen's lives and sustainability of the city. Moreover, the collected data can be vertically integrated reducing data silo to be used to develop new applications (Krishnamachari et al. 2018). Hence, data from smart devices, meters, sensors, social media, etc. have become one of the important assets for enterprise that provide digital services to citizens, while presenting data owners/seller and data consumer/buyers with prospects for new insights into achieving profitable business (Bajoudah and Missier 2021).

The monetization of data has become an on-trend business, not only for enterprises that deploy smart devices in cities, but also for citizens. With this revolution any residents in cities with metering devices and smart sensors can sell his/her information such as energy usage records, energy production records, energy supplier information or even energy usage patterns in order to earn revenue as value can be gotten from the streams of data (Bajoudah and Missier 2021). But one of the challenges in developing smart cities lies in the management of complexities associated with different platforms and technologies. With different components and smart devices, the data generated in smart cities could get fragmented resulting to silos (Xu et al. 2019). Respectively, there is need for research that improve and supports data distribution, establish trustworthiness of data source via ratings to enable data consumers to choose data sources based on an established scale for economic, social, environmental, and technological aspects of urban environment (Money and Cohen 2015).

2.2 Background of data marketplaces in smart cities

The rise of digitalization in cities has resulted to massive amounts of data are being generated from different infrastructures having thousands of smart meters, sensor devices, digital platforms, and legacy systems which provide comparatively static data (Jeong et al. 2020; Jnr et al. 2021). Data of importance are produced by citizens, industrial vendors, or the government (Fricker and Maksimov 2017). The use of data aids decision making by offering significant improvements to many areas of the economy, including in platforms deployed in the areas of mobility, transportation, energy, health, crowd-sensing, etc. Thus, service providers in cities such as Uber are progressively operating and deploying data driven systems for optimizing their operations using integrated data from different sources, and occasionally, even combining data from different areas. Data are now regarded as a key factor, as pointed out by Hatamian (2021) where the authors stated that data are a useful resources in enterprises and cities. Most enterprises operating in cities often require data that they cannot collect on their own, thus they utilize commercial data marketplaces (Azcoitia et al. 2020). A data marketplace is an evolution of data repositories as seen in Fig. 2.

Figure 2 depicts the evolution of data marketplaces in smart cities. Data marketplaces is the next evolution in data management which provides a scalable solution for cities in moving forward to achieve data driven innovations. It is an interactive digital platform that facilitates the interactions between data owner and data consumer (Serrano 2021). Data marketplaces are explicitly designed to address issues associated with data lake. By refining and extending data lake which is a storage repository that retains a huge volume of raw data in its native format till it is needed for data analytics applications. Thus, in smart cities data marketplaces enables stakeholders such as developer, enterprises, citizens, urban administration to know the location of their data and track and monitor its usage. By

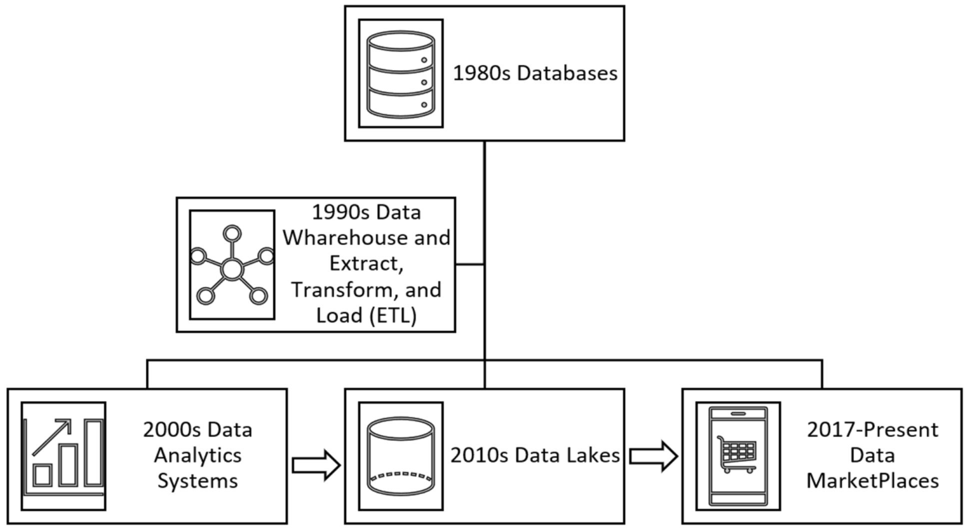


Fig. 2 Evolution of data marketplaces adapted from (Early Adopter Research 2019)

knowing what data are being utilized, the municipality can make more informed decisions regarding which data are to be improved and how to governance data across the city (Early Adopter Research 2019).

In a data marketplace data owners register their datasets along with standardized descriptions or metadata on what the data are about, domain information, cost, and all other pertinent information (Mišura and Žagar 2016; Fricker and Maksimov 2017; Serrano 2021). It allows the selling and purchasing of data from individuals, companies, or governments (Mišura and Žagar 2016). Also, data consumers or buyers can query the data marketplace and indicate what data type they need (Mišura and Žagar 2016). In data marketplaces, data owner, providers or sellers advertise their data to pursue potential data consumer or buyers. Data marketplace provides a platform that promotes data sharing in smart cities as these systems collect data from different types of devices installed across the city (Hatamian 2021). Data from the data marketplace can be used by different applications to address a variety of purposes in smart cities such as model creation, data analytics, visualizations, and city dashboards, etc. Service providers within the city can offer value added services based on available data assets to support municipality administration in the embarking on data-driven projects (Gindl 2020).

A data marketplace could help cities share relevant data and further help partners in the municipalities develop better-improved digital services. Data marketplaces are required to act as a neutral intermediary which allow interested data sellers to upload and advertise their data products to be sold digitally (Hatamian 2021). Interested data buyers search and select the proposed data they want. In traditional or centralized data marketplaces the marketplace mediator gets incentive for aiding the trade and, in most cases, for hosting the data (Banerjee and Ruj 2018). The data marketplace platform queries the appropriate data and provides data consumers access to the dataset under terms that both data owners and data consumers find acceptable for a specific fee. The data are usually published and purchased in different file textual format such as excel tables, full databases (Mišura and Žagar 2016; Banerjee and Ruj 2018), or may be accessible as online streams of new data or static archives (Fricker and Maksimov 2017). Different forms of data access can be presented, e.g., via application programming interface (API) for accessing whole data repositories or through subscriptions (Fricker and Maksimov 2017; Hatamian 2021).

In urban environment new systems are being developed for different domains resulting to data silo and fragmentation. Thus, cities systems become less interoperable resulting to vendor lock-in as it becomes challenging to re-use collected data (Jeong et al. 2020), on weather, traffic, noise, vehicles, and air quality (Bajoudah and Missier 2021). Furthermore, Fig. 3 depicts an overview of data marketplaces in smart cities based on the categories and perspectives. As seen in Fig. 3 data marketplaces differs with regards to the participants type (either business-to-business (B2B), customer-to-business (C2B), or any), the geographic scope of the platform (which includes local for a particular city, regional, or global across cities), from a technical viewpoint regarding the deployed architecture (centralized or decentralized) (Gindl 2020), where the role of decentralized data marketplaces in making cities smarter are discussed in Sect. 2.3. According to Stahl et al. (2016) data marketplaces are categorized in three different areas which comprises of *independent, private, and consortium*. A private data marketplace has either a single data supplier or single data buyer, such that it offers many-to-one or one-to-many relations to the data buyers/suppliers. Consortium data marketplaces are unions of

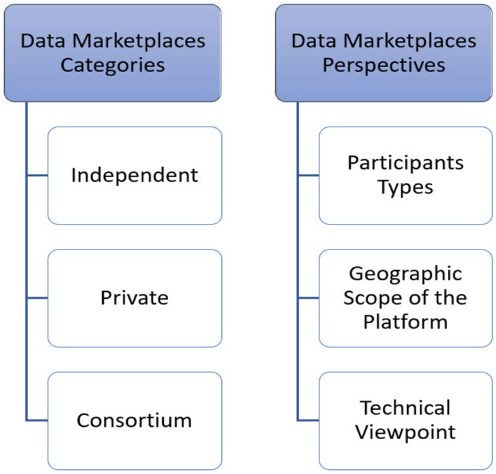


Fig. 3 Overview of data marketplaces in smart cities

both data suppliers and data buyers, or hybrid of both. The independent data marketplace operates as an intermediary for data buyers and data suppliers, for instance in providing infrastructures and platform for effective data search, procurement, settling, and transfer of datasets.

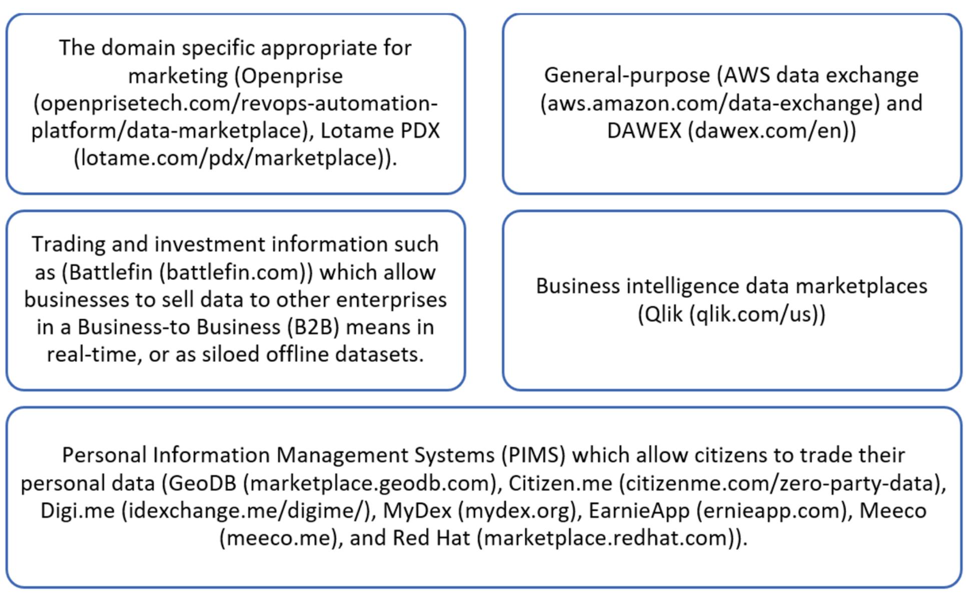
The independent data marketplace allows for many-to-many relation or data sales between independent stakeholders in the data marketplace platform (Gindl 2020). As stated by Azcoitia et al. (2020) there are five different types of data marketplaces applicable in smart cities as seen in Fig. 4. Accordingly, Fig. 4 shows the different types of data marketplaces applicable in smart cities which comprises of personal information management systems, trading and investment information, business intelligence, domain specific data marketplaces, and lastly general-purpose data marketplaces.

2.3 Decentralized data marketplaces in making cities smarter

The adoption of digital technologies in the past decade has brought about notable disruption across sectors. One of such popularly emerging or digital technologies is DLT. A DLT is mainly a distributed records or ledger of transactions (Lawrenz et al. 2019). Once data are entered in the DLT, the information is immutable. Other benefits of DLT includes improved traceability, more transparency, efficiency and lower costs, and enhanced security (Lawrenz et al. 2019). In brief, DLT is a decentralized database such as blockchain which was initially used to carryout commercial transactions among independent participants without trusting on a centralized authority, such as government agencies or banks. DLT such as blockchain technology has become a popular area since the advent of cryptocurrency such as Bitcoin in 2008. Basically, blockchain is a public ledger centered on consensus protocols to provide a verifiable, append chained based data structure of transactions (Lawrenz et al. 2019; Xu et al. 2019). Due to the decentralized architecture DLT allows data to be updated and stored distributivity. Also, in blockchain data transactions are authorized by miners and stored in the time-stamped blocks, where individual blocks are identified based on a cryptographic hash and connected to preceding blocks in a sequential order. Within a blockchain network, a consensus protocol is instantiated on many distributed nodes called miners to preserve the sanctity of the data stored on the blocks (Xu et al. 2019; Anthony Jnr and Abbas Petersen 2021).

The decentralized or peer-to-peer distributed network-based data marketplace offers a much fairer and trustworthy platform as compared to the conventional centralized system where all data transactions are managed by a

Fig. 4 Types of data marketplaces applicable in smart cities



central server. Also, to ensure resilience of the data marketplace there is no single point of failure thus there is less risk of the entire digital platform failing. In a decentralized data marketplace, all contributing peers within the network communicate with similar privileges since it is a distributed network (Park et al. 2018). In a data marketplace data provider uploads the data to advertise these datasets to data consumers who access the data marketplace using the peer-to-peer distributed network. The datasets uploaded by the data provider comprises of raw encrypted data saved on the cloud server. Further, the data provider only discloses the hash value or metadata to the data consumer and provides a key such as a decryption access key that gives access to the cloud server's raw data once the data transaction is verified. However, the limitation of this decentralized data marketplace approach is that data consumers cannot completely trust the shared dataset (Park et al. 2018).

Additionally, most DLT such as blockchain Ethereum uses smart contracts. These smart contracts are self-executing scripts deployed within blockchain. When a predefined condition within smart contract the among participating parties such as data provider and data consumer are met then the individuals involved in a contractual agreement can autonomously make payments based on the contract in a transparent way. Smart contract supports transparency of data transaction to all involved individuals and is suitable as a medium for data transactions (Lawrenz et al. 2019). DLT aids the implementation of decentralized application (DApp) which is developed in the form of a smart contract and executed on a DLT network to deploy predefined agreements and algorithms without depending on a third-party involvement. Therefore, smart contracts and Dapp can be employed to provide a decentralized solution to address challenges in data marketplaces (Xu and Chen 2021).

Presently, there are few available DLT based data marketplaces which have been commercialized for use cases in smart cities (Abbas et al. 2021), such as IOTA data marketplace (IOTA Foundation 2022), Ocean Protocol (Oceanprotocol 2022), Data Exchange Platform (Dawex 2022), Data Intelligence Hub (Data Intelligence Hub 2020), and Advaneo (Advaneo 2022). These decentralized data marketplaces employed decentralized architecture and are mostly implemented employing smart contracts. DLT such as IOTA tangle, Tendermint, HyperLedger Fabric, Ethereum, etc. has shown a great potential to digitally transform business and the society. Thus, DLT is an ideal technology to be employed for data marketplace based on its decentralized architecture which facilitates distributed transactions among all business members in a trustless environment with built-in aid for micropayments in smart cities.

2.4 Related works of data marketplaces in smart cities

One of the means by which a city can become smarter and sustainable is to develop the local economy around the sharing and use of data from open data and other data sources that can be utilized in applications to enhance the lives of its citizens. Due to the potential of DLT in driving digital transformation in the society blockchain based platforms have been designed in prior literature to facilitate trading of IoT data in smart cities. Other interesting studies such as Bajoudah and Missier (2021) examined the latency of trading transactions in a Ethereum brokered based IoT data marketplace. Findings from their research described the technical deployment and evaluation of a developed marketplace model. The marketplace improves the trading of streaming data, by offering data for sale and exchanging trade agreements to automate trade fulfilment and obligations. Meijers et al. (2021) designed a cost-effective blockchain-based IoT data marketplaces based on a credit invariant. The study presented a trustless data trading platform that decreases the risk of fraud and the number of transactions completed on chain. A credit mechanism was also created to further lower the incurred fees.

Additionally, Pomp et al. (2021) designed a semantic driven data marketplace to aid easy sharing of data within a smart city. Findings from their study presented a data marketplace that supports various stakeholders (citizens enterprises, and public institutions), to easily deliver data that can mainly contribute to the further actualization of smart cities. The marketplace is grounded on the principles of semantic data management, where data contributors annotate their added dataset with semantic models. Gindl (2020) delivered a report on data marketplaces aligned to interoperability solutions to explicitly support the exchange of data with third parties. The report offered technical solutions and best practices for interoperability by proving recommendations on concrete guidance on how to deploy, implement, and integrate interoperability solutions for use in data marketplaces. Na et al. (2020) explored how to improve the reliability of IoT data marketplaces via security validation of IoT devices. The authors assessed the security of IoT data marketplaces and investigated different types of vulnerabilities that exist in IoT data marketplaces grounded on the well-known STRIDE model, and further offered a security assessment and certification framework for IoT data marketplaces to support device owners to assess the security vulnerabilities of their IoT devices. Most importantly, the presented approach helps application developers make informed decision when purchasing and utilizing data from a data marketplace.

Another study by Lawrenz et al. (2019) employed blockchain as an approach to achieve data marketplaces. The study is aligned to Recycling 4.0 which focused on improving sustainable recycling method through data exchange. The study identified data quality, data integrity, and secure platform as some major issues that impacts the operability of data marketplace. Xu et al. (2019) suggested a blockchain facilitated secure microservices for achieving a decentralized data marketplace. Using a permissioned blockchain network a microservices based security method is deployed to secure data exchange and payment among contributors within the marketplace. The suggested approach offers a decentralized, auditable, and scalable data exchanges for data marketplace. Banerjee and Ruj (2018) investigated the design and challenges associated with developing data marketplaces. The study proposed a blockchain supported data marketplace system that helps to address issues related to adhering to regulations, efficiency, fairness, privacy, and security. Krishnamachari et al. (2018) designed an IoT based marketplace for smart communities termed the I3 concept. The designed approach comprises of edge devices, I3 platform, data analytics, IoT application with different stakeholders (device/data owners, marketplace operators, data broker, application developers, and end users). The authors further implemented a simple proof of concept prototype using MQTT pub-sub broker on the backend supported by a Python Django-based web marketplace frontend using a MySQL database.

Ramachandran et al. (2018) explored decentralized data marketplace within the context of smart cities. The author investigated how a decentralized data marketplace can be developed using blockchain and other DLTs. The study considered the potential advantages of a decentralized architecture in data marketplaces, identified different components that a decentralized marketplace should have, and illustrates how these could be possibly integrated into a complete solution. The authors also presented a simple smart contract decentralized registry where data products can be published by data owners and be used by potential data buyers. Park et al. (2018) implemented a smart contract-oriented review application for an IoT data marketplace. The implemented system can verify the reputation of the data traded or a data owner in a P2P data marketplace. The system uses Ethereum smart contracts deployed on the P2P network and provides a flexibility as compared to traditional server-client review systems which have many shortcomings, such as server administrator's malicious behavior or security vulnerability. Mišura and Žagar (2016) investigated data marketplaces for IoT. The study developed a model of a market, analyzed the query performance, estimated profits of IoT device owners and utility derived by data consumers. Findings from the study stated that data markets in IoT are a feasible method of delivering generated data measurements.

Prior work in the literature have focused on the development of centralized data marketplaces for smart cities (Mišura and Žagar 2016; Krishnamachari et al. 2018; Gindl 2020; Na et al. 2020) and another study on semantics (Pomp et al. 2021). A few other have explored decentralized data marketplaces by employing blockchain (Banerjee and Ruj 2018; Ramachandran et al. 2018; Lawrenz et al. 2019; Xu et al. 2019; Meijers et al. 2021), Ethereum (Park et al. 2018; Bajoudah and Missier 2021), etc. This current study examines how a decentralized data marketplace could be designed using IOTA tangle and MQTT protocol as a broker which has not been utilized in the literature. This study is different form prior studies as the author(s) further considers the potential benefits of a decentralized brokered approach, identify different components required within such decentralized data marketplace, and depicts how this approach could be theoretically and practically integrated into a comprehensive data solution. This decentralized IOTA brokered based marketplace aids citizens, enterprises, and municipalities to easily share and sell their data with all interested stakeholders.

3 Method

In this study, IOTA tangle enabled by MQTT is employed to design a distributed ecosystem to facilitate trust and fairness among data owners and data sellers in data marketplace within smart cities. Each of these technologies are discussed below;

3.1 IOTA technology for data marketplaces in smart cities

IOTA is a public DLT which was initially release live in July 2016 to encrypted ledgers (Akhtar et al. 2021), that applies an underlying data structure referred to as directed acyclic graph (DAG) storing structure referred to as Tangle, for saving transactions. IOTA provides a fully decentralized restricted, private, and public architectural platform. Similar to the Wei (in Ethereum) or Satoshi (in Bitcoin), IOTA's native cryptocurrency termed "*iota*", is represented as "*i*" which is the smallest unit within the IOTA network (Akhtar et al. 2021). As an alternative to blockchain, the DAG only needs each transaction entered within the tangle to approve two prior transactions. Any unapproved data transaction is known as a "*tip*" and the more transactions that authorize other transaction the more trustworthy the system has with the new transaction (Gagnon and Stephen, 2018).

Previously, micropayments were never possible in existing DLT platforms as transaction fees were much higher than the business transactions, but IOTA employs fee-less environment which supports micropayments (Akhtar et al., 2021). IOTA offers a miner-free system where all IOTA tokens have been established (Akhtar et al. 2021; Anthony Jnr 2022). Thus, IOTA employs a special type of PoW consensus, in which each node that adds a transaction to the distributed ledger must authenticate two other transactions. Thus, IOTA offers decentralization, any node can simply join the distributed network or contribute to the consensus mechanism (Akhtar et al., 2021). Such an authentication mechanism allows IOTA to provide free transactions as compared to other PoW-based systems such as Ethereum and BitCoin (Ramachandran et al. 2019). IOTA is mainly designed for industrial sectors which uses IoT and machine-to-machine (M2M) which offers a direct communication between devices in smart cities utilizing any communications channel, including wireless and wired.

IOTA provides zero fees for miners and supports micro-transaction system via a secure data transmission protocol (Xu and Chen 2021). IOTA supports peer-to-peer data transactions without any centralized bodies that collects data from different sensors in smart cities via APIs (Serrano 2021). IOTA tangle is preferred as compared to other DLT such as blockchain as it offers functionalities that are needed to establish a M2M micropayment system (Popov 2018), which is suitable in data marketplaces. On the contrary, IOTA also introduce its own issues. For example, it may be challenging for IOTA tangle to directly monitor data flow of via MQTT connections. This eventually makes fair enforcement of trade contract difficult. Also, the interoperability of IOTA tangle with legacy systems and external centralized digital platforms may be an issue (Anthony Jnr 2021a).

3.2 Applicability of publish-subscribe communication and MQTT protocol

Over the decades enterprise applications such as IBM WebSphere and Microsoft Azure Service Bus employs the publish-subscribe communication mechanism to exchange data between data producers and consumers. This messaging protocol separates the subscribers and publishers in synchrony, space, and time. Publish-subscribe communication mechanism is a substitute to point-to-point request-reply and synchronous communication models. The main advantages of the publish-subscribe communication mechanism is its ability to facilitate loosely coupled message exchanges between consumers and producers (Ramachandran et al. 2019). The publish-subscribe messaging protocol has proven itself as a dominant messaging communication approach for IoT and urban systems. By detaching publishers (data sources) from the subscribers (data sinks) and clients (either subscribers or publishers) synchronously. Most distributed platforms rely on a messaging mechanism for collaborating and coordinating with other digital systems within the network.

A few messaging protocols proposed for distributed interactions includes remote procedure call (RPC), constrained application protocol (CoAP), and message queueing telemetry transport (MQTT) (Ramachandran et al. 2019). Clients within the MQTT network can publish their messages to the broker based on self-defined topic. Clients may also subscribe to get updates on current topics. Overall, the MQTT broker broadcasts live updates and available clients subscribed to a particular topic (Ramachandran et al. 2019). In smart cities MQTT and CoAP are mostly employed as they are suitable for IoT and digital applications. Although CoAP is less used due to its request and reply's communication approach, whereas MQTT only requires one request to be sent to a message broker, after which data from the producers are redirected to data subscribers by the message broker (Ramachandran et al. 2019). MQTT is lightweight in terms of resource intensiveness which makes it possible to connect to millions of IoT devices within smart cities where data can be collected to be shared in data marketplaces.

In this study, the Message Queuing Telemetry Transport (MQTT) which is a messaging protocol based on the OASIS standard is employed (Ramachandran et al. 2019). The publish-subscribe communication approach usually comprises of three components which are the publisher, topic (broker), and subscriber as seen in Fig. 5 (Anthony Jnr et al. 2020). Figure 5 illustrates an example of publish-subscribe communication based on MQTT for measuring movement data from foot counter IoT device to server and clients. The publishers are responsible to send data to the broker using the percepts of the topic. The topic usually refers to the metadata, which defines information regarding the data in a string format. A topic can have several levels (Ramachandran et al. 2019). For instance, the data produced by a foot counter IoT device deployed at the city center within a smart city can have its topic described as \city_center_region1\street123\citizen_movement_counter. The municipality administration and other stakeholders who utilizes this data can receive data from the foot counter IoT device by subscribing to the \city_center_region1\street123\citizen_movement_counter topic.

As seen in Fig. 5 the foot counter IoT device publishes updates about the number of movements recorded at specific time intervals within a particular location. It publishes the number of movement data to the MQTT broker, who then broadcasts it to subscribed clients which are the municipality administration and other stakeholders. Although the publish-subscribe messaging is resource-efficient, scalable, and lightweight. It depends on a central broker for effective data communication among publishers and subscribers. Failure of the broker may affect the subscribers and publishers. Additionally, the conventional publish-subscribe brokers

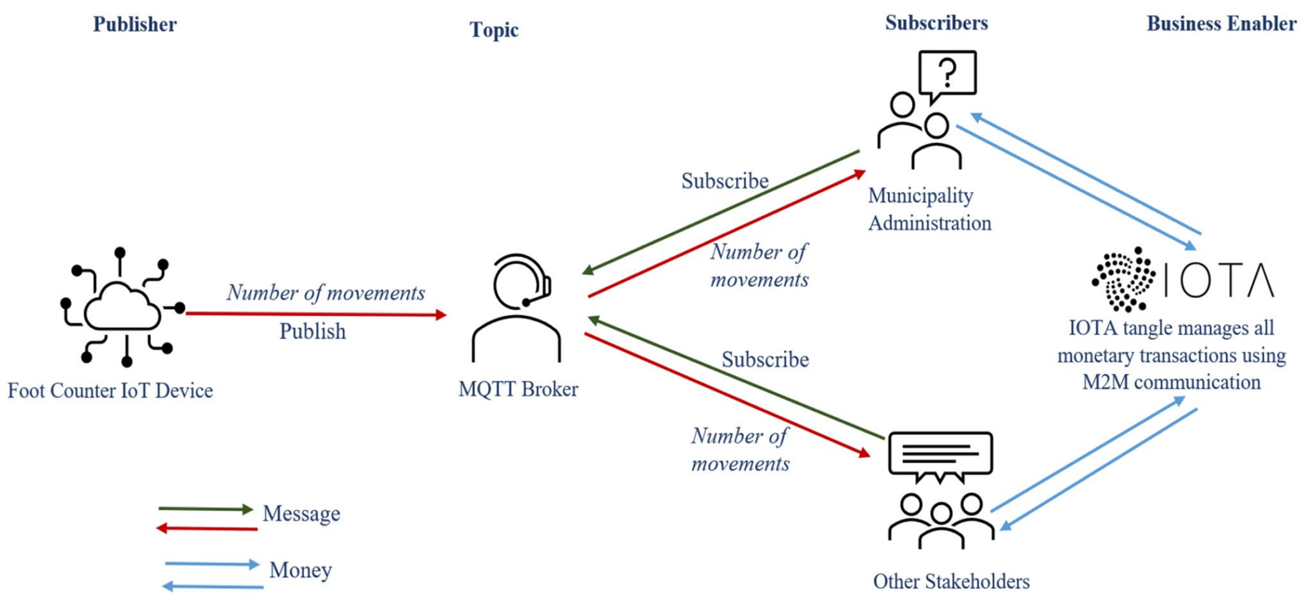


Fig. 5 Publish-subscribe approach based MQTT adapted from Anthony Jnr et al. (2020)

do not provide accurate guarantees involving the planned delivery of messages to the subscribers, though some application support ordered delivery for a single broker setting. Also, MQTT broker is pluggable with DLT platforms such as IOTA, HyperLedger Fabric, Ethereum, and Tendermint (Ramachandran et al. 2019). As shown in Fig. 5 IOTA tangle is included as a *business enabler* to aid the "transfer of money" for sale of dataset among the municipality and other stakeholders within the city interested in owning or getting access to data needed to create value added services (Bokolo 2022).

3.2.1 Significance of mosquitto eclipse for MQTT

In smart cities the implementations of publish-subscribe MQTT protocol can be achieved using brokers such as Mosquitto (Eclipse Mosquitto 2022). MQTT has previously been deployed in the areas of smart homes, automotive, logistics, manufacturing, etc. to developed interoperable prototype integrated with the Eclipse Mosquitto MQTT broker and the Eclipse Paho MQTT Python Client as a client library (Gindl, 2020). The Eclipse Mosquitto was mainly developed for usage as a single instance, but it can provide support for integrating multiple connected brokers to work together via the bridging functionality which efficiently duplicates all communicated messages at every broker, allowing subscribers and publishers to connect to any instance. But this method of data distribution does not assure the same listing of messages at every broker. Also, in deploying Eclipse Mosquitto for data communication or transactions there is also an inherent assumption of the platforms being deployed by a single entity, hence there are no trust concerns, although the data immutability is not provided (Ramachandran et al. 2019).

3.3 Designed ecosystem for data marketplaces

The ecosystem for data marketplaces is designed based on IOTA tangle and MQTT protocol as seen in Fig. 6.

To develop data driven services in smart cities there is need to be develop decentralized data marketplaces that allow datasets from different sources to be merged and analyzed, processed, and acted upon to support the deployment of diverse applications. As pointed out by Ramachandran et al. (2018) a seller of data in a city with one or more static, dynamic, or even streaming data products (that could be data generated from an IoT device such as the foot counter (as seen in Fig. 5) counting movement of citizens or anything from building occupancy to air quality, and data buyers interested in getting this data are the main participants in the decentralized data marketplace.

3.4 Data collection and analysis

In this study, exploratory research was employed with the help of a qualitative method by collecting data from interview and secondary data from the literature. The interviews provide an unstructured way to provide information that allows a deep knowledge of phenomena, problems, or situations (Delgado-de Miguel et al. 2019). The interview conversations provide much deeper information than other data collection procedures, such as surveys, because they allow the researchers to clarify doubts and acquire a good number of details. In connection

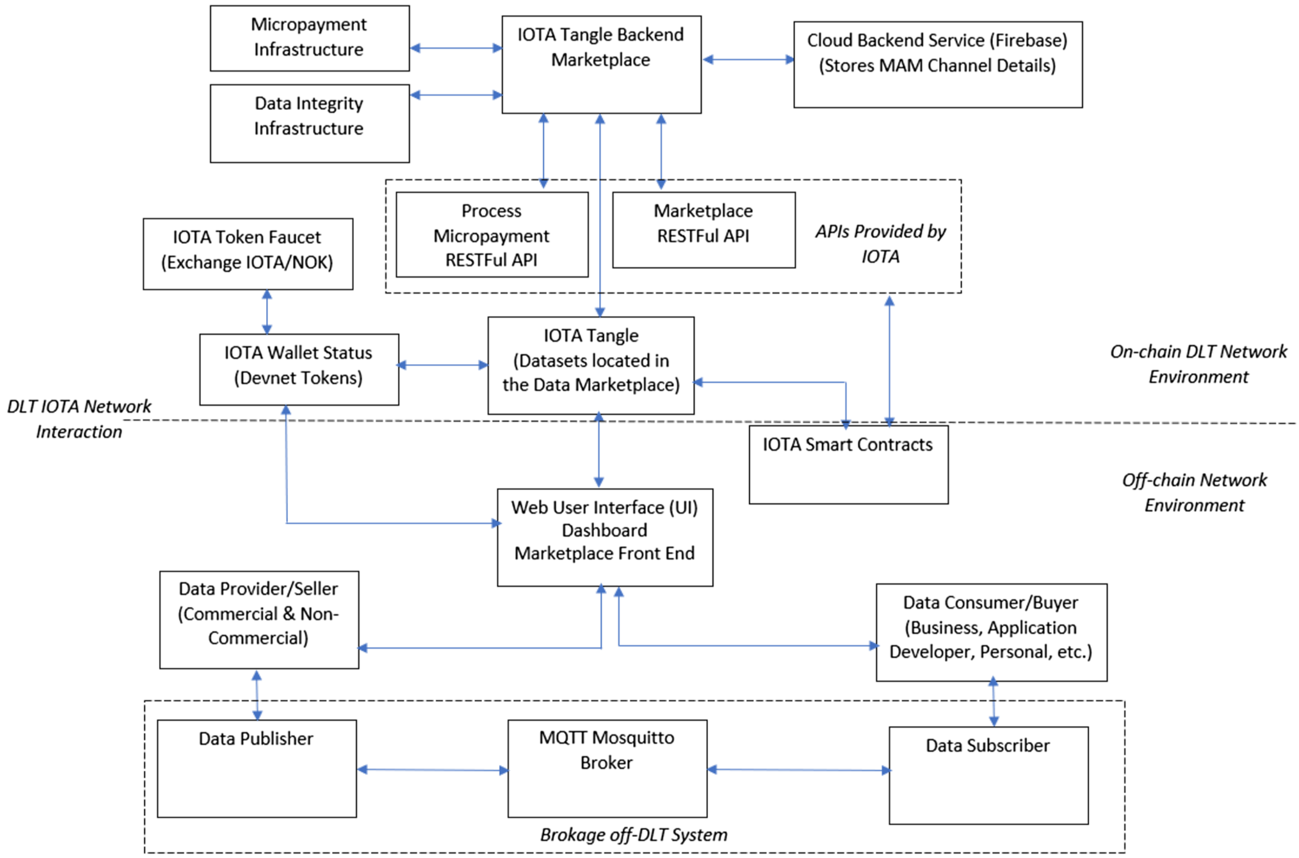


Fig. 6 Designed ecosystem for data marketplaces

to the + CityxChange smart city project (https://cityxchange.eu/) where IOTA was involved in ensuring data integrity, immutability, and auditable in energy trading and electric mobility as a service in smart city. Semi-structured interviews were conducted with some partners in IOTA which comprises of (head of telecommunication and infrastructure development, with up to 20 years of experience in designing digital solutions for IoT. Another informant is the project manager with almost 4 years of experience involved in developing emergent solutions using IOTA tangle in the areas of IoT, data integrity, verification, micro-financial transactions, data sharing, tokenization of assets, data access control, self-sovereign identity, and smart contracts. The next respondent is a senior technical analyst with about 3 years' experience with how IOTA tangle can improve mobility services in smart cities and has contributed to the ideation, design and implementing of digital services that make data usable in the most efficient way.

Besides other partners in the + CityxChange smart city project involved in the use of IOTA tangle were involved in providing feedback related to use of IOTA tangle for providing data driven services such as supporting energy bidding and trading and micropayment in electric mobility as a service. The interviews focused on the participant`s knowledge of the business processes of IOTA tangle in providing data driven services in smart cities and how decisions were made in deployment of IOTA in helping cities to be sustainable. A manual analysis technique was used to code the qualitative data, with a descriptive approach to determine the informants' assumptions of IOTA as a suitable DLT to provide to support data marketplace. Thus, responses from the interview sessions were manually recorded while additional evidence form document provided by IOTA. Also, technical reports gotten from IOTA website (https://www.iota.org/solutions/smart-city), were used to get more insights. In addition, core themes related to the designed ecosystem for data marketplace were extracted. Although the findings were limited as only few small numbers of informants were involved in the data collection. These preliminary findings from the literature and interview provide evidence to the context of data marketplaces for continued research in this area (Sutherland and Hovorka 2014).

4 Findings

Primary data from the interview session and secondary data from the literature were included to model the ecosystem for data marketplace. Figure 6 shows the ecosystem which allows diverse data owners to contribute to the data economy by selling datasets, while different data consumers such as application developers can connect via a web User Interface (UI) to buy and obtain one or more datasets that are useful to their application. The ecosystem for data marketplace as shown in Fig. 6 is categorized based on two layers similar to prior study (Bajoudah and Missier 2021). Overall, the ecosystem for data marketplace involves an integration of different sections (on-chain DLT network environment, off-chain network environment and the brokage off-DLT system), that are deployed separately and then linked together to achieve a decentralized brokered enabled ecosystem for data marketplace in smart cities where all the data marketplace functionalities are provided. The on-chain DLT network environment represents the use of IOTA tangle and the IOTA smart contract, which aids to enforce and automate the legal obligations of data trades. Therefore, once the data buyers and data seller agree on the sale of a dataset the transaction will be carried out by IOTA tangle and IOTA smart contract. Whereas the off-chain network environment mainly comprises of the web user marketplace dashboard frontend which is responsible to make calls and interact with IOTA tangle in the form of sending trade transactions, reading metadata, and making payment settlement in local currency or in IOTA token.

Also, the ecosystem facilitates the exchange of datasets which is executed off-chain facilitated by the brokerage platform (MQTT Mosquitto Broker), from transactions that occur between data producers and data consumers. Besides, in the ecosystem employs IOTA uses Representational State Transfer RESTful APIs (process micropayment API and marketplace API) based on a JavaScript object notation (JSON) data structure to facilitate micropayment and orchestration of the data marketplace. One of the stakeholders in the ecosystem are the data consumer who are set of data buyers interested in purchasing datasets from the data marketplace. Another stakeholder is the data provider are potential sellers who are willing to sell their datasets in the data marketplace. The data sellers can enter meta data description within a JSON template which provides information description of the data. The metadata information comprises of data description, including the seller`s identification, type of data, IP address, price, etc. in JSON format stored in a distributed file storage (Ramachandran et al. 2018).

Additionally, only the data hash or metadata values are disclosed in the data marketplace platform due to limited storage space and security concerns the actual datasets are not presented directly in the data marketplace platform (Park et al. 2018). After the transaction has been confirmed, the data sellers provide the data buyers with a key to access the dataset stored in the IOTA tangle with the payments happening via IOTA tokens from the IOTA wallets and the resulting transactions are recorded in the IOTA tangle. After the end of the data transaction, both the sellers and the buyers can rate each other. The rating functionality is accessible to the data buyers to rate and possible comment on the dataset being traded. This helps to verify the reputation of the data sellers to check whether the data being traded are trustworthy based on the previously submitted reviews. This information is stored in IOTA tangle to ensure the reviews cannot be altered and the integrity and immutability is ensured.

A summarized findings as seen in Fig. 6 is presented in Table 1

Table 1 provides a summary of the findings for the practicality of the ecosystem components to enable sale of data marketplaces.

4.1 Applicability of the designed ecosystem for trading dataset

Within the designed ecosystem to give the data sellers an easy way to provide dataset, the data marketplace will offer a wizard which guide data sellers in entering some textual description of the dataset (Ramachandran et al. 2018; Pomp et al. 2021), into a data product description template provided in the data marketplace frontend where data sellers specifies some information of the dataset (Ramachandran et al. 2018), known as the metadata which describes the data product. This will help to specify how the dataset is made available (either commercial or open data once, and or continuously periodically), through which data source the dataset is available upon purchase (e.g., via MQTT stream or FTP server) and in which data format (e.g., JSON, CSV, XML, or others) the dataset is provided (Pomp et al. 2021). The metadata which describes data sources will be provided in a structured manner by each data sellers can organize hierarchically based on the template that is used by the data sellers to provide information of the dataset to be solved (Ramachandran et al. 2018).

To enable data transfer and payments, once a data buyer is interested in a dataset, he/she can proceed to select and proceed to payment preferably via the cryptocurrency-based payment channel (IOTA token) or through local currency. This can be implemented via the micropayments in IOTA payment channel or by employing the streaming data payment protocol (SDPP) (IOTA Foundation 2020), which is a protocol that is appropriate for real-time data with micropayments which utilizes a TCP client–server communication for

Table 1 Summarized findings for the usefulness of the ecosystem components

Components	Description
Micropayment infrastructure	This technology is used by IOTA to support the processing of micro payment for digital services such as mobilities payment for transport solution or energy trading in smart cities
Data integrity infrastructure	This is a technology deployed by IOTA tangle backend to support data sharing, guaranting data integrity, immutability, and auditable. It also ensures identity verification, access control, data integrity, participants rating, and instantiating privacy policies
IOTA tangle backend	This is the main control center of the IOTA tangle. The backend controls input and output of data with other applications via APIs provided by either IOTA or other third-party systems
Cloud backend service (Firebase)	Stores Masked Authenticated Messaging (MAM) channel details of IOTA tangle. Also, ensures logged data are private and only available for parties with the MAM channel key within the tangle
APIs provided by IOTA	Provides access to data to support sale of datasets within the decentralized data marketplaces platform. This API provided a bridge as connectors between IOTA tangle backend and the data marketplace platform
IOTA token faucet	Aids the creation of an IOTA wallet for data sellers and data buyers. Then, set dataset price, loads IOTA wallet, and supports the exchange of IOTA digital currency (IOTA tokens), to "NOK" Norwegian local currency from the IOTA wallet to a connected user account. Also, it helps data sellers to set up of dataset price in NOK/IOTA. It further supports the exchange of a given number of NOK to IOTA tokens and transfer to the IOTA wallet. Moreover, it helps for data sellers and data buyers to check wallet balance. Enabling holding and transfer of IOTA tokens to another IOTA wallet by processing direct peer-to-peer payments between data sellers and data buyers using IOTA tokens
IOTA wallet status	This mainly stores the status of IOTA tokens used as a legal tender within the data marketplace to carryout payment of datasets by data buyers. The pay can be made wither in the local currency (e.g. NOK) or in IOTA tokens
IOTA tangle	Among other things records information of all datasets available in the data marketplace and provides information about the agreed sale of datasets. Aids to send IOTA transactions to and from IOTA wallets for data sellers and data consumers. Moreover, provides secure data verification and integrity, locally creates IOTA transactions (set up buying and selling prices), enables a direct payment between the two involved parties' data consumer and data producer. Also, it requests payments in IOTA tokens directly from data consumers and data producers, and requests payment to be sent to the IOTA smart contract once data sale is fulfilled after payment is confirmed
IOTA smart contracts	IOTA smart contract refers to an immutable state machines which connects to the IOTA Tangle. As a state machine it comprises of data information such as input conditions, account balances, and consequences over time to update a state transition within the tangle. Although the IOTA smart contract`s program code and state are both unchangeable since they are stored within the tangle. But the state can be incrementally updated by appending new data transactions to the tangle, which provides a verifiable audit trail for each state transitions (Iota Beginners Guide 2022) The IOTA smart contract is also responsible to ensure payment requests between assets are shared using the IOTA ledger to ensure auditability. Besides, it logs sold datasets to IOTA tangle via APIs (process micro payment API and marketplace API), provided by IOTA and refund appropriate amount from appropriate data sellers when cancelation is made by data buyers
Web UI interface	Supports data producers and data consumers to interact and conduct data trading in the front end supported by the IOTA tangle backend. Also, provide a user interactive dashboard where data buyer and data sellers can view some information of dataset, payment history, and other profile information
Data provider/seller	Provides both commercial and non-commercial (open data) datasets from different sources such as citizen data, traffic data, sensor data, social media data, etc. These datasets are advertised in the data market platform front end for potential data buyer or data consumers
Data consumer/buyer	These are the end users who are interested in acquiring data from the data marketplace platform. The data are used to provide data driven services in smart cities
Brokered off-DLT system	Comprises of the data publisher, data subscribers, and MQTT Mosquitto broker. This part of the ecosystem is responsible to receives demand and offers for datasets, matches data demands and offers from potential data consumers, and optimally matches request to possible datasets. Also, send information to clients such as application developers, municipality administrators, citizens, enterprises, etc. on available dataset based on their subscribed topics

data transfer, and a DLT channel to save the key transaction records such as sales receipts and invoices (Ramachandran et al. 2018). Lastly, the data buyer can rate the data quality after purchase of the dataset. This enables data consumers to evaluate the quality of the data product provided by a given data sellers. Also, the data sellers can also rate the data buyers as well. The data sellers could assess data buyer's reliability of paying for data delivered and the data buyer could rate the timeliness and quality of data received. These information regarding the ratings would then be made accessible for future potential data consumers and data providers as recommendations for other prospective data buyers (Ramachandran et al. 2018).

4.2 Challenges and recommendations

4.2.1 Challenges in decentralized data marketplaces

There are many challenges and requirements associated with data marketplace, such as security data integrity which verifies that the dataset has not been altered at any point and guaranteeing data immutability (Lawrenz et al. 2019). But these issues are addressed in a decentralized data marketplace such as an IOTA based data marketplace, which are unfortunately faced with other issues such as data quality, data storage, privacy, etc. These issues are discussed below:

a. Privacy threats

One of the most discussed topics in decentralized data marketplaces is privacy. IoT devices in smart cities typically collect and produce personal data that can be used to obtain private information about citizens. Privacy is the claim of persons, groups, or organizations to define for themselves how, when, and to what point information about them is shared to others. This implies that users can trade their private data in data markets if they know exactly what data they are trading and to who (Mišura and Žagar 2016). But there are privacy related threats within decentralized data marketplaces since everyone who access the distributed ledger can view some information of the data sellers and data buyers. Thus, privacy and copyright protection are still a challenge (Hatamian 2021). Without an appropriate privacy protection, this could lead to significant privacy violations. One of the most considerable requirements from the General Data Protection Regulation (GDPR) is the “right to rectification” and “right to be forgotten”.

The right to be forgotten suggests that the data provider has the right to the erasure of private data regarding him/her without undue delay. Furthermore, rectification implies that the data providers should be able to update and correct inaccurate data concerning him/her. Accordingly, in data marketplaces the most crucial challenge as related to privacy is protecting private sensitive information while maintaining the decentralization of the dataset published. All personal data must be Pseudonymized as a solution but how to accomplish this within the DLT may also be an issue if hashes cannot be used. Also, DLT based platforms do not generally conform to GDPR because of the decentralized nature of DLTs (Hatamian 2021). Moreover, most countries have different privacy regulation and laws hence trade transactions related to data sharing may be complicated (Banerjee and Ruj 2018).

b. Governance process

Governance entails policies and strategies within the data marketplaces which is important in this study as this involves the control mechanisms involved for self-regulation for fairness (Linkov et al. 2018), trust, and transparency for data sharing (Abbas et al. 2021). Governance comprises of the administration processes by certain actors (e.g., data marketplace operators), deployed mechanisms, such as norms or power (Linkov et al. 2015). The governance processes involved in managing any decentralized based platform is always not clear as there is no central authority in these systems (Anthony Jnr 2022). But in the case of the decentralized data marketplace there are certain actors who manage the platform such as the data marketplace operators via various mechanisms, raising the need to know the norms or power distribution in data marketplaces.

c. Storage capabilities

Storage infrastructures form the vital part of data marketplaces. Today, several data marketplaces are utilizing centralized storage systems by adopting cloud storage services such as Google Drive, MongoDB, and MySQL. Data storage in cloud has emerged as a method to resolve the exploitation of storage space to meet the increased production of data in smart cities. However, these storage solutions not completely secured (Hatamian 2021). While storage capabilities are the underlying part of a distributed data marketplaces, achieving continuous storage within DLTs is still an issue. This is because data stored in DLTs are immutable and cannot be changed but new update for the data can only be appended to the initial version. This can result to lack of storage in future.

d. Data quality

One of the issues faced in data marketplace is ensuring data quality. Data quality ensures that data are generally tailored for use by data consumers (Adams et al. 2019; Lawrenz et al. 2019;). Additionally, data buyers should be able to verify the quality of data before purchase. Although the metadata can be assessed. This may not be enough to evaluate the quality of the data as data sellers may provide fake data (Lawrenz et al. 2019). However, getting a suitable method for data quality assessment before data buyer purchase a dataset is still challenging regardless of all the current solutions (Mišura and Žagar 2016).

e. Data integrity

Also, a mechanism is required to ensure the integrity of all datasets added by the data sellers. Data integrity is the guarantees that the datasets are consistent and has not been distorted through their lifecycle. Ensuring data integrity of data source assess that the data are legitimate and has not been altered (Lawrenz et al. 2019). Ensuring trust is another issue faces generally in data markets. This is because malicious users could purchase data from data owners, and then re-sell the data to other data consumers at a lower price to make profit thereby depriving the true data owner of this benefits. This issue is similar to copyright infringement and mechanisms should be imposed to prevent this from occurring (Mišura and Žagar 2016).

f. Social effects

Social implications of data marketplace relate to the impacts of decentralized data marketplace for the benefit of the society at large, such as ethical issues related to data sharing and other consequences of data trading for cultural, political, social, and economic viewpoint (Abbas et al. 2021).

g. Pricing mechanism

An interesting issue is how to set a fair price for commercial dataset. In almost all commercial data marketplaces, pricing of datasets is decided by the data sellers and the data buyers need to agree (Cooke and Golub 2020). Data sellers employ different pricing scheme such as open data/free, flat fee tariff, pay-per-use, freemium (Gindl 2020), or set a fixed price, or let data buyers bid for datasets, or even a hybrid combination of two methods (Azcoitia et al. 2020).

h. Adhering to fair principles

Most applications in smart cities generate data from different locations by several stakeholders. These data should be findable, accessible, interoperable, and reusable based on the FAIR principles. While the FAIR principles have basically become a main component of data management, its applicability for data sharing within smart cities is limited. While most open data are often already incompatible across urban systems due to data collected from heterogeneous data sources. To support a better data sharing in smart cities, data marketplace platforms should allow the publishing and sharing of data from different stakeholders with the respect to the findability, accessibility, interoperability, and reusability (FAIR) principles (Pomp et al. 2021).

Table 2 presents a summary of the possible risks faced in decentralized data marketplaces and risk mitigation initiatives that can be employed to address the identified risks.

4.2.2 Recommendations for decentralized data marketplaces

There should be fairness in data marketplaces therefore before performing the trade of data, the final price of the dataset needs to be agreed upon by the partners. Also, there should be transparency, privacy, and security for the dataset which is being traded. Therefore, the dataset must not be disclosed to any other individuals not involved in the sale. Public log should be employed to determine the ownership of data and to safeguard that only legitimate parties gain access to the data itself. Transparency in terms of pricing should be deployed (Banerjee and Ruj 2018). Additionally, different types of regulations such as health insurance portability and accountability act (HIPAA) related to health dataset should be adhere to. Any breach of these regulations may reduce the trust among the partners involved in trading and buying of data.

Although, most regulation such as GDPR can be complicated to follow because of data trade across geographical regions which may conflict laws between countries. To set price for the data a computational pricing

Table 2 Summarized risks and associated mitigations within decentralized data marketplaces

#	Possible risks	Risk mitigations
1	Privacy issues if hashes cannot be used to ensure security	Pseudonymized personal data within the data marketplaces
2	The governance processes involved in managing any decentralized based platform is always not clear	Enable data logs and versioning which records data ownership and access for data in the marketplaces
3	Storage solutions may not be completely secured	Employ off-chain cloud storage services such as Google Drive, MongoDB, and MySQL which has inhouse hash algorithms
4	Ensuring data quality is another issue faced in data marketplace	Use metadata to assessed data quality made available within the data marketplace
5	Ensuring the integrity of all datasets added by the data sellers	Provide data provenance to ensure complete data lifecycle
6	Ethical issues related to data sharing is not well resolved	Employ DLT which complies to privacy regulations such as HIPAA, GDPR, etc. to reduce unethical issues
7	Difficulty to set a transparent and fair price for commercial dataset	Employ smart pricing algorithm, hybrid pricing or biding schemes for datasets to ensure price transparency
8	Incompatible across urban systems due to data collected from heterogeneous data sources	Comply to the FAIR principles to support the publishing and sharing of data from different stakeholders

algorithm, smart pricing algorithm based on Stackelberg game theory, or machine learning-based algorithms that helps for price determination mechanisms can be employed. These algorithms can be based on training data or pre-trained models. The smart pricing algorithm was previously applied in blockchain-based data marketplaces to support data brokers increase their revenue (Abbas et al. 2021). Lastly, the efficiency of the system should be improved as this will influence a widespread adoption the data marketplace. An efficient decentralized data marketplace should have an improved speed as compared to other conventional data marketplaces (Banerjee and Ruj 2018).

5 Discussion and implications

5.1 Discussion

With the advent of data marketplaces, cities around the world can develop data driven applications which utilizes different data sources (Na et al. 2020). In smart cities it is necessary to develop an ecosystem that supports the interoperable sharing of data from different sources in a fair and trustworthy approach. This is because data are a strategic asset that can help municipalities secure and maintain sustainable competitiveness. Cities can benefit from available data in developing better services driven by large scale data. Consequently, data marketplaces have emerged as a platform to enable transparent data transactions between data producers, owners or sellers and data consumers or buyers. Multiple data producer can make their data available either as commercial or non-commercial for the application developers and other stakeholders via data marketplaces (Na et al. 2020). Data producers sell their data for economic benefits or incentive or to build trust, and data consumers acquire and use these datasets (Park et al. 2018; Xu and Chen 2021). Therefore, data marketplaces as well as Intelligent IoT Integrator, Ocean Protocol, and IOTA data marketplace (IOTA Foundation 2022), are being developed to improve the use of data to improve cities in to becoming smart communities and cities (Na et al. 2020). But the success of data marketplaces depends on the trust, fairness, and reputation of the data products (Mišura and Žagar 2016).

Accordingly, this study integrates IOTA tangle, IOTA smart contract and MQTT broker to increase data sellers' and data buyers' trust for data marketplaces platforms. To realize the "ecosystem for data marketplaces", a publish-subscribe broker which provides supports for topic-based communication is employed. Therefore, in this study the publish-subscribe broker method is employed by the Eclipse Mosquitto as the broker which coordinates the communication between clients (partners) (data provider/seller as publishers and data consumer/buyer as subscribers) and the IOTA tangle platform. As stated by Banks and Gupta (2014) the broker instance is connects to a consensus node orchestrated by the DLT client library offered by the underlying DLT platform which is IOTA tangle in the case of this study (see Fig. 6), which interacts with the MQTT Mosquitto Broker REST APIs. Likewise, the functionalities of the broker are deployed on top of MQTT using predefined APIs. As illustrated in Fig. 6 IOTA is employed in the designed ecosystem which address immutability issue faced by the Eclipse Mosquitto broker. IOTA maintains a decentralized distributed database ledger called as tangle which stores transactions permanently ensuring immutability for all registered traded data transactions managed by the IOTA smart contract. All transactions are transparently accessed by cryptographic hash functions thus, reducing malicious attack (Park et al. 2018).

The designed ecosystem for data marketplaces provides the data to application developers, municipality administration, and other users for implementing digital applications that horizontally integrate silo data from different platforms and applications. In centralized data marketplace, a mediator is required to support data exchange from different domains and origins, integrating different data types and stipulating pricing mechanisms (Abbas et al. 2021). Centralized data marketplaces are vulnerable to a single point of failure since the central control can influence the marketplace processes or initiate malicious policies to boost profits while misleading data sellers and/or data buyers (Avyukt et al. 2021). To resolve the setbacks faced by server-client model, the decentralized approach has emerged, in which all nodes contribute equally to the distributed network (Park et al. 2018). Accordingly, the decentralized data marketplace is proposed in this research. The decentralized data marketplaces, which leverage DLT and MQTT broker to involve multiple parties in the business process, enhancing transparency and trust. Findings from this study depicts how IOTA tangle is employed to ensure the integrity, immutability, and reliability of transactions within the data marketplace. This finding is analogous with results from prior study (Park et al. 2018), where a blockchain platform Ethereum was applied effectively as it offers a cryptocurrency function and but also provide smart contracts that allow other distributed functionality to develop a Dapp that manages reviews of data sold within a data marketplace.

The designed ecosystem for data marketplaces in this study supports sharing of non-commercial or open data similar to findings from the literature (Jeong et al. 2020), where the authors use open data portals to provide static data (e.g., logs and statistics) which were seldomly updated and provided as manual based file formats. The findings from this study suggest that APIs (process micropayment API and marketplace API), are used to support trading of data in smart cities. This result is in line with results from the SynchroniCity project which implemented a data marketplace grounded on business APIs ecosystem from FIWARE in the Korean smart cities open data portals facilitated by oneM2M standard APIs (Jeong et al. 2020). This finding is also in line with results from Xu et al. (2019) where data was extracted via APIs and stored in a cloud server. Xu et al. (2019) validated a data marketplace model within Newcastle urban observatory smart city project using data which has been extracted by open protocol RESTful APIs based on a JSON data structure. The authors executed microservices by exposing REST APIs to receive service requests. Additionally, finding from this study is similar to results from prior study Ramachandran et al. (2019), where the MQTT pub-sub broker using blockchain technology was deployed on an immutable distributed ledger to guarantee data persistence. Thus, each data source that is added to the marketplace platforms are annotated with a semantic standard based on a knowledge graph's underlying ontology to ensure that all data related to smart city applications development adhere to FAIR principles.

5.2 Implications for research and practice

5.2.1 Research implications

In smart cities IoT devices generate potentially valuable which have huge impact to the society and are being used in large sectors of the economy, such as automotives, health care, manufacturing, urban environment, and other domains. However, as many of the IoT devices operate as isolated and closed systems. These issues can be limited by giving data producers or owners appropriate economic incentives to trade their data through efficient, trustworthy, fair, and secure marketplaces (Meijers et al. 2021). Data marketplaces are being developed in the context of smart cities. Data marketplaces are anticipated to play a critical role in tomorrow's data economy towards economic sustainability in smart cities. A data marketplace which is multi-sided digital platform matches data providers and data buyers by facilitating a data sharing economy. Data sharing in smart cities via data marketplaces may contribute to overall economic development by stimulating data driven innovation, enhancing the competitiveness of enterprises, and opens new job opportunities (Abbas et al. 2021). Commercialization of data marketplaces supports the creation of new digital products and services as it beneficial for businesses that do not have exclusive access to data required to improve their revenue.

Theoretically this study adds to the literature by providing an extensive overview of current practice of data marketplace by identifying the components needed to support a decentralized data marketplace in smart cities and how these components could be deployed without trusted third parties or centralized servers. The finding identifies some of the key issues inherent in achieving a decentralized data marketplace in smart cities. IOTA tangle, MQTT protocol, and IOTA smart contract together are promising technologies to provide a solution to enable a secured, fair, and trustworthy data sharing. For academics, findings from this study offer insights into the issues that impacts the actualization of a decentralize data marketplace. Also, recommendations are provided to improve the development of decentralize data marketplaces in general.

5.2.2 Practical implications

The focus of this study is the design of components for data provision and consumption that accelerate sharing and trading of dataset in a decentralized data marketplace. To this end an ecosystem for data marketplace is designed driven by IOTA tangle and MQTT protocol. As suggested by Krishnamachari et al. (2018) this study design an economically self-sustaining business ecosystem around data that provides incentives for citizens, municipalities, and companies. It provides a level of trust between trade participants by providing data advertising options driven by MQTT, offering the ability for negotiate a data trade agreement between data sellers and data buyers. It aids in fulfilling data trade agreement transactions with the deployment of DLT represented by the IOTA tangle and IOTA smart contract which manages payment in IOTA tokens or local currency and rating for trade participants. More importantly, this study goes beyond the basic description of data marketplaces but also presents practical insights regarding technical designs of an ecosystem for a decentralized broker-based data marketplace that could potentially instill trust and fairness among data owners and data sellers when trading data in smart cities.

Findings from this study offer guidance of how a decentralize data marketplace can be developed for smart cities. The decentralized ecosystem for data marketplace guarantees the integrity of the data and makes sales transactions fair, such as writing or submitting trade reviews, transparent to everyone to avert malicious behavior. Practitioners interested in data marketplace deployment in smart cities can reflect on these findings. As the designed ecosystem for data marketplace provides a viable business model. The finding provides a state-of-the-art to broaden the practitioners understanding on decentralized data marketplaces. Practitioners can also reflect on the designed ecosystem for data marketplace to explore potential value added services for stakeholders in smart cities. This research provides the theoretical and practical roadmap for future research toward the implementation of an IOTA brokered based data marketplace. This study explores how a decentralized data marketplace could be designed and deployed using distributed ledger technologies and publish-subscribe communication.

6 Conclusion

This research contributes to the understanding of decentralized data markets in smart cities by illustrating the role of DLT as an emerging technology which contributes towards sustainable development of the city. Findings from the literature suggest that data are needed to have an interoperable, replicable, and sustainable smart city. Therefore, this study designs an ecosystem that supports the actualization of a decentralized brokered enabled data marketplace for smarter cities towards a data driven society. IOTA tangle is deployed as the DLT and MQTT protocol is deployed as the broker to facilitate data exchanges and data trading. The ecosystem aims to potentially instill trust and fairness among data owners and data sellers when trading data in smart cities. Exploratory research was employed with the help of a qualitative method by collecting data from interview and secondary sources to provide insight on the usefulness of IOTA tangle to support data sharing guaranting data integrity, immutability, and auditable. Insights from this research are expected to inform municipalities administrators, empowered application developers and business operating within the city on the potential of data in making cities smarter.

The findings also present the usefulness of data in making cities smarter, background of data marketplaces in smart cities, significance of decentralized data marketplaces in making cities smarter and related works related to data marketplaces in smart cities. Finally, open challenges that impacts decentralized data marketplaces and recommendations are provided to improve the deployment of data marketplaces. In future work, it is envisioned that a data marketplace prototype based on the designed ecosystem will be implemented using python programming preferably Django for the frontend connected to MQTT Mosquitto and IOTA tangle to validate the feasibility of the ecosystem. Additionally, to successfully design and commercialize decentralized data marketplaces, it is crucial to identify the complete stakeholders involved in data marketplaces to understand the distribution of incentives within the ecosystem. Therefore, future research can also focus on studying the governance of decentralized data marketplace to examine the roles of stakeholders within data marketplaces. Additionally, the implemented data marketplace prototype based on MQTT Mosquitto and IOTA tangle will be further examined using survey questionnaire and interview data to assess the usefulness and effectiveness of the tool.

Acknowledgements The author is grateful to Norwegian University of Science and Technology, NTNU and the +CityxChange smart city project (https://cityxchange.eu/) which is funded by the European Union's Horizon 2020 Research and Innovation Programme under the Smart Cities and Communities topic with Grant Agreement No. 824260. The interview data employed in this article were part of the +CityxChange smart city project analyzed by the author during his postdoctoral fellowship at NTNU.

Author contributions The author contributed in drafting the manuscript.

Funding Open access funding provided by Ostfold University College. This study did not receive any funding.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

References

Abbas AE, Agahari W, van de Ven M, Zuiderwijk A, de Reuver M (2021) Business data sharing through data marketplaces: a systematic literature review. J Theor Appl Electron Commer Res 16(7):3321–3339

Adams S, Greenspan S, Velez-Rojas M, Mankovski S, Beling PA (2019) Data-driven simulation for energy consumption estimation in a smart home. Environ Syst Decis 39(3):281–294

Advaneo (2022), ADVANEO data marketplace. https://www.advaneo-datamarketplace.de/en/ Accessed online on 5 Apr 2022

Akhtar MM, Rizvi DR, Ahad MA, Kanhere SS, Amjad M, Coviello G (2021) Efficient data communication using distributed ledger technology and iota-enabled internet of things for a future machine-to-machine economy. Sensors 21(13):4354

Alvsvåg R, Bokolo A Jr, Petersen SA (2022) The role of a data marketplace for innovation and value-added services in smart and sustainable cities. Innovations for community services: 22nd international conference, I4CS 2022, Delft, The Netherlands, June 13–15, 2022, Proceedings. Springer International Publishing, Cham, pp 215–230

Anthony Jnr B (2021a) Distributed ledger and decentralised technology adoption for smart digital transition in collaborative enterprise. Enterp Info Syst. https://doi.org/10.1080/17517575.2021.1989494

Anthony Jnr B (2021b) Managing digital transformation of smart cities through enterprise architecture–a review and research agenda. Enterp Info Syst 15(3):299–331

Anthony Jnr B (2021c) Exploring data driven initiatives for smart city development: empirical evidence from techno-stakeholders' perspective. Urban Res Pract. https://doi.org/10.1080/17535069.2020.1869816

Anthony Jnr B (2022) Toward a collaborative governance model for distributed ledger technology adoption in organizations. Environ Syst Decis 42(2):276–294

Anthony Jnr B, Abbas Petersen S (2021) Examining the digitalisation of virtual enterprises amidst the COVID-19 pandemic: a systematic and meta-analysis. Enterp Info Syst 15(5):617–650

Anthony Jnr B, Abbas Petersen S, Ahlers D, Krogstie J (2020) API deployment for big data management towards sustainable energy prosumption in smart cities-a layered architecture perspective. Int J Sustain Energ 39(3):263–289

Avyukt A, Ramachandran G, Krishnamachari B (2021) A decentralized review system for data marketplaces. IEEE international conference on blockchain and cryptocurrency (ICBC). IEEE, pp 1–9

Azcoitia SA, Paraschiv M, Laoutaris N (2020) Computing the relative value of spatio-temporal data in wholesale and retail data marketplaces. Preprint at https://arxiv.org/abs/2002.11193.

Bajoudah S, Missier P (2021) Latency of trading transactions in brokered IoT data marketplace in Ethereum. 2021 IEEE smartworld, ubiquitous intelligence & computing, advanced & trusted computing, scalable computing & communications, internet of people and smart city innovation (SmartWorld/SCALCOM/UIC/ATC/IOP/SCI). IEEE, pp 254–263

Banerjee P, Ruj S (2018) Blockchain enabled data marketplace--design and challenges. Preprint at https://arxiv.org/abs/1811.11462.

Banks A, Gupta R (2014) MQTT Version 3.1. 1. *OASIS standard*, *29*, 89.

Barroso M, Laborda J (2022) Digital transformation and the emergence of the Fintech sector: systematic literature review. Digital Bus 2:100028

Bokolo AJ (2022) Exploring interoperability of distributed ledger and decentralized technology adoption in virtual enterprises. IseB 20(4):685–718

Cooke R, Golub A (2020) Market-based methods for monetizing uncertainty reduction. Environ Syst Decis 40(1):3–13

Cosgrave E, Arbuthnot K, Tryfonas T (2013) Living labs, innovation districts and information marketplaces: a systems approach for smart cities. Proc Comput Sci 16:668–677

Data Intelligence Hub (2020) Data intelligence hub. https://dih.telekom.net/en/ Accessed online on 5 Apr 2022

Dawex (2022) Data exchange, unleash the value of your data. https://www.dawex.com/en/ Accessed online on 5 Apr 2022

Delgado-de Miguel JF, Buil-Lopez Menchero T, Esteban-Navarro MA, Garcia-Madurga MA (2019) Proximity trade and urban sustainability: small retailers' expectations towards local online marketplaces. Sustainability 11(24):7199

Early Adopter Research (2019) The transformative impact of the data marketplace WHITE PAPER. https://www.qlik.com/us/-/media/files/resource-library/global-us/register/whitepapers/wp.transformative-impact-of-the-data-marketplace-en.pdf Accessed online on 8 Apr 2022

Fricker SA, Maksimov YV (2017) Pricing of data products in data marketplaces. International conference of software business. Springer, Cham, pp 49–66

Gindl S (2020) Data marketplaces with interoperability solutions I. Nature 31:12

Hatamian M (2021) Technological barriers of (non) blockchain enabled IoT data marketplaces. 2021 25th International computer science and engineering conference (ICSEC). IEEE, pp 39–44

Iota Beginners Guide (2022) Iota beginners guide-IOTA SMART CONTRACTS. https://iota-beginners-guide.com/products/iota-smart-contracts/ Accessed online on 12 Apr 2022

IOTA Foundation (2020) Understanding SDPP on IOTA. https://ecosystem.iota.org/tutorials/understanding-sdpp-on-iota Accessed online on 13 Apr 2022

IOTA Foundation (2022) IOTA data marketplace. https://data.iota.org/ Accessed online on 27 Mar 2022

Jeong S, Kim S, Kim J (2020) City data hub: Implementation of standard-based smart city data platform for interoperability. Sensors 20(23):7000

Jnr BA, Petersen SA (2022) Using an extended technology acceptance model to predict enterprise architecture adoption in making cities smarter. Environ Syst Decis. https://doi.org/10.1007/s10669-022-09867-x

Jnr BA, Petersen SA, Ahlers D, Krogstie J (2020) Big data driven multi-tier architecture for electric mobility as a service in smart cities: a design science approach. Int J Energy Sect Manage 14(5):1023–1047

Jnr BA, Petersen SA, Helfert M, Ahlers D, Krogstie J (2021) Modeling pervasive platforms and digital services for smart urban transformation using an enterprise architecture framework. Inf Technol People 34(4):1285–1312

Kieti J, Waema TM, Ndemo EB, Omwansa TK, Baumüller H (2021) Sources of value creation in aggregator platforms for digital services in agriculture-insights from likely users in Kenya. Digital Bus 1(2):100007

Krishnamachari B, Power J, Kim SH, Shahabi C (2018) I3: An iot marketplace for smart communities. Proceedings of the 16th annual international conference on mobile systems, applications, and services. IEEE, pp 498–499

Lawrenz S, Sharma P, Rausch A (2019) Blockchain technology as an approach for data marketplaces. Proceedings of the 2019 international conference on blockchain technology. ACM, pp 55–59

Legenvre H, Hameri AP, Golini R (2022) Ecosystems and supply chains: how do they differ and relate. Digital Bus 2(2):100029

Linkov I, Larkin S, Lambert JH (2015) Concepts and approaches to resilience in a variety of governance and regulatory domains. Environ Syst Decis 35(2):183–184

Linkov I, Trump BD, Anklam E, Berube D, Boisseasu P, Cummings C, Vermeire T (2018) Comparative, collaborative, and integrative risk governance for emerging technologies. Environ Syst Decis 38(2):170–176

Marikyan D, Papagiannidis S, Rana OF, Ranjan R (2022) Blockchain: a business model innovation analysis. Digital Bus 2:100033

Meijers J, Putra GD, Kotsialou G, Kanhere SS, Veneris A (2021) Cost-effective blockchain-based iot data marketplaces with a credit invariant. 2021 IEEE international conference on blockchain and cryptocurrency (ICBC). IEEE, pp 1–9

Mišura K, Žagar M (2016) Data marketplace for Internet of Things. 2016 International conference on smart systems and technologies (SST). IEEE, pp 255–260

Money WH, Cohen S (2015) Developing a marketplace for smart cities foundational services with policy and trust. Int J Comput Sci Theory Appl 3(1):1–12

Mosquitto E (2022). Eclipse Mosquitto™ An open source MQTT broker. https://mosquitto.org/ Accessed online on 27 March 2022

Na Y, Joo Y, Lee H, Zhao X, Sajan KK, Ramachandran G, Krishnamachari B (2020) Enhancing the reliability of IoT data marketplaces through security validation of IoT devices. 2020 16th international conference on distributed computing in sensor systems (DCOSS). DCOSS, pp 265–272

Oceanprotocol (2022) Tools for the Web3 data economy-ocean market. https://market.oceanprotocol.com/ Accessed online on 5 Apr 2022

Park JS, Youn TY, Kim HB, Rhee KH, Shin SU (2018) Smart contract-based review system for an IoT data marketplace. Sensors 18(10):3577

Pomp A, Paulus A, Burgdorf A, Meisen T (2021) A semantic data marketplace for easy data sharing within a smart city. Proceeding

of the ACM international conference on information & knowledge management. ACM, pp 4774–4778
Popov S (2018) The tangle, iota whitepaper. IOTA, Tech. Rep.[Online]. https://iota.org/IOTA_Whitepaper.pdf
Ramachandran GS, Radhakrishnan R, Krishnamachari B (2018) Towards a decentralized data marketplace for smart cities. 2018 IEEE international smart cities conference (ISC2). IEEE, pp 1–8
Ramachandran GS, Wright KL, Zheng L, Navaney P, Naveed M, Krishnamachari B, Dhaliwal J (2019) Trinity: A byzantine fault-tolerant distributed publish-subscribe system with immutable blockchain-based persistence. 2019 IEEE International Conference on Blockchain and Cryptocurrency (ICBC). IEEE, pp 227–235
Serrano W (2021) Validation and verification of data marketplaces. IFIP international conference on artificial intelligence applications and innovations. Springer, pp 780–792
Stahl F, Schomm F, Vossen G, Vomfell L (2016) A classification framework for data marketplaces. Vietnam J Comput Sci 3(3):137–143
Sutherland D, Hovorka D (2014) Enterprise architecture as a contributor to sustainability objectives. Proceeding of the European conference on information systems (ECIS). ECIS
Vaninsky A (2021) Multiobjective restructuring aimed at green economic growth. Environ Syst Decis 41(1):110–130
Xu R, Chen Y (2021) Fed-ddm: a federated ledgers based framework for hierarchical decentralized data marketplaces. 2021 international conference on computer communications and networks (ICCCN). IEEE, pp 1–8
Xu R, Ramachandran GS, Chen Y, Krishnamachari B (2019) Blendsm-ddm: Blockchain-enabled secure microservices for decentralized data marketplaces. 2019 IEEE international smart cities conference (ISC2). IEEE, pp 14–17