Data Access and Usage Control for Data Spaces

Ramon da Gama Cordeiro[1,*], Marcela Tuler de Oliveira[1]

[1]*Delft University of Technology, Mekelweg 5, 2628 CD Delft, The Netherlands*

Abstract

Data sharing has become increasingly prevalent, and effective access and usage control mechanisms have become fundamental to protect sensitive information and ensure trust among partners. Policy enforcement to ensure safe access and usage of the data is imperative in a collaborative environment. Data space requires fine-grained control over the data so as not to poison the reliability of the data-sharing environment. Additionally, it is necessary to ensure the auditability and trustworthiness of data usage. However, some solutions have been developed to address those problems, and many apply a centralised system design approach. The challenge of centralised architecture is to ensure auditability, which is not built into this architecture and can reduce the system's trustworthiness. Distributed Ledger Technology emerges as an alternative to ensure reliability, auditability, and security. It can be applied in data spaces to improve control over the data with auto-executed policies within smart contracts. The paper presents a research proposal using a decentralised approach to address data access and usage control in a data space environment. The proposal is designed to use distributed ledger technology with smart contracts to enforce control policies.

Keywords

Access Control, Usage Control, Distributed Ledger Technology (DLT), Smart Contracts, Attribute-Based Access Control, Data Sovereignty, Data Spaces

1. Introduction

Digital systems are facing a challenge with access and usage control, especially because there has been a rise in the amount of data collected and processed in the last few years, opening up very new applications of data-driven solutions. Moreover, the advent of data leverage approaches such as big data and AI brings concerns about how, when and who uses the data, and why [1]. Those concerns have been materialised in some actions and regulations like GDPR [2], Data Act[1], and AI Act [3].

The GDPR [2] defines some actors related to data, *data subject* is who the data relates to, and this data contains an identifier or identifiable information about this person. *Data controller* [4] is a person or institution that determines the purpose and how the data will be processed. Multiple institutions can act collectively as *Joint Controllers* [5]. The data processor is responsible for processing the data according to the data controller definitions [6].

Data sovereignty is the concept that a natural or legal person decides sovereignly using their data as an asset [7]. Additionally, a data processing agreement is a legal contract between the data controller and the data processor defining the rights and obligations of each part [8].

The legal initiatives were made to guarantee data sovereignty and auditability, which helps to highlight the discussion about the sovereignty of the data and recover the importance of data sovereignty as an important element to obeying the laws, preserving fundamental rights, freedoms and the right to protect personal data [2]. Auditability is a technical way to ensure that the laws and jurisdiction rules are being applied. Additionally, auditability enables data controllers to verify the data usage for which they are responsible.

Although legal initiatives, such as GDPR, are important, there are still gaps that laws cannot reach without technical solutions to address the problem. Some of those gaps in the technical field are related to access control and usage control, especially regarding sensitive data. Nowadays, access control solutions work with data controllers, requiring sovereignty over their data and a third party to actually

DLT2025: 7th Distributed Ledger Technology Workshop, June, 12-14 2025 - Pizzo, Italy

[*] Corresponding author.

r.ramoncordeiro@tudelft.nl (R. d. G. Cordeiro); m.tulerdeoliveira@tudelft.nl (M. T. d. Oliveira)
0009-0003-0383-8685 (R. d. G. Cordeiro); 0000-0001-5739-5590 (M. T. d. Oliveira)

manage the access to the data. This approach is not enough for institutions with sensitive data that want to be part of a collaborative environment. Furthermore, the absence of sovereignty over the data by the o raises questions about the trustworthiness of digital systems. Once they share the data, they lose control of further usage, and if the original purpose of data sharing is maintained. With data today, predicting behaviour, market trends, institutional decisions, and more is possible. Thus, the data must be used according to data controller definitions and context, preserving rights, privacy and auditability. Then, data controllers must take control over the policies on how, when, and why their data is or will be used [9].

Most access control mechanisms are based on centralized architectures, where a single institution or company, such as a cloud provider, holds complete control over access decisions. Usually, a centralised system has limited access to data processing logs and how transparent the logs and the whole system are. An alternative to this challenge and the limitations of centralised architecture can be to apply decentralised solutions with transparent and auditable logs, such as Distributed Ledger Technology (DLT). In this way, DLT enables the use of data in data spaces with sovereignty, auditability, and the possibility of privacy-preserving [10] [7].

This paper defines a research direction for addressing data access and usage control in data spaces by applying a decentralised and auditable manner of access and data usage, applying policy enforcement, and using smart contracts. One of the big questions is how to ensure data usage following the compliance of GDPR, ensuring traceability of the usage, giving control of the data to the data controller and revoking access when misbehaviour occurs. Moreover, usually, other proposed mechanisms do not enforce data usage, do not ensure the data control to the data controller at all or do not track the data usage transparently to the controller. We are trying to go beyond access control but proposing an approach to usage control as well. This paper proposes an approach to address these questions.

The rest of this paper is organised as follows. Section 2 focused on the background. Section 3 presents the related works. In section 4, it is presented research problem and motivation. In section 5 it is presented the proposed approach to address the challenge. In section 6 we describe expected contributions. In section 7 we present the research plan. Finaly in section 8 we present the conclusion.

2. Background

2.1. Data space and Simpl Open

Data space is a framework that supports data sharing within a data collaborative environment. It is a way to store and share data using relationships and relevance between the data [10]. The data is aggregated based on subject-related data and their relationship. Data space has component subjects, services and controllable datasets. At the same time, controllable datasets refer to the amount of data under the responsibility or stewardship of a data controller. Additionally, the controllable dataset has objects (content data) and their relationship (metadata). Services are the subject's actions or features according to its data and needs. The management and data control are subject-oriented in that services such as filtering, storing and classifying are made according to the subject's definitions [11].

The nature of data spaces, which allows data controllers to take over their data, enables data spaces to be used in a trusted environment for sharing data. There is a high demand for exchanging data with trustworthiness, obeying regulatory compliance, and protecting sensitive data [10]. Thus, data space enables cooperation between institutions and integration of secure data exchange, such as healthcare data.

There are some initiatives about data spaces for sharing data and promoting innovation and development of the data-driven applications as part of the European strategy for data applied by the European Commission [9]. Simpl is an open-source middleware platform supporting data access and interoperability between data spaces [12]. The main goal of this project is to create a Common European Data Space ecosystem that makes it possible to share data safely and compliantly with regulations. The Simpl architecture comprises three components: SIMPL-Open, SIMPL-Lab and SIMPL-Live. The SIMPL-Open is an open-source software stack that enables data spaces and federated solutions, in general, to share data. SIMPL-Lab is an assessment environment for the data spaces to evaluate their capabilities, such as their interoperability level. Simpl-Lab is also used to experiment with SIMPL features before running them in the execution engine. Finally, the SIMPL-Live is the execution engine which runs SIMPL-Open for a set of data spaces.

The Simpl initiative aims to establish a unified European data market governed by standard policies and regulations. In a data space, participants, data controllers and data processors can share data securely, transparently, and with mutual trust. Simpl is tailored for the public and private sectors. However, the initiative needs more development regarding policy definitions and enforcement mechanisms.

2.2. Access and Usage control

Access control can be defined as a set of rules which permit or deny access to specific data or services. An authority can be a data controller or service provider that sets the rules for data access.

Usage control refers to a set of policies that rule data use after granting access. The usage control extends policies to the lifecycle of data usage. Moreover, data usage can enforce obligations, limitations of use, and purpose, and retain control after the data controller share their data. Unlike access control, the data controller cannot define policies of having control after data is shared [13].

Simpl Open recommends Role-based Access Control (RBAC). RBAC is a mechanism that permits or denies access to resources based on the roles of the users. RBAC is widely adopted because it is simple to manage permissions. The permissions are set to a role which can be linked to a group of users, and these users have access to resources based on the role's permissions. Moreover, if a data manager revokes the permissions to a role, all users with that role have their permissions revoked [14]. The main idea of RBAC is hierarchical and responsibility access, in a way that users within the same level of responsibilities and hierarchy likely have the same level of information access. RBAC is composed of four components: roles, users, permissions and resources. Roles are a collection of permissions associated with a function within an organisation. Users are personnel who were assigned roles to access certain resources. Permissions are specifications of logical rules that are translated into actions of a user on data. Resources are assets, files, and data that are accessed under policies and restrictions.

Attributed-based Access Control (ABAC) is a dynamic and extensible access control mechanism that grants or denies access based on a set of attributes associated with subject, resource, action attributes, and environment attributes [15]. In addition, ABAC ensures fine-grained access control and more flexibility in enforcing policies based on contextual attributes. ABAC is composed by some elements that works together and each has a role to ensure fine-grained access control. ABAC relies on a set of key components — namely, the Policy Administration Point (PAP), Policy Decision Point (PDP), Policy Enforcement Point (PEP), Policy Information Point (PIP), and Context Handler (CH) — each of which plays a specific role in ensuring fine-grained, dynamic access and usage control. The PAP stores and manages a set of policies that are used to evaluate access requests. The PDP is where decisions are made when a request arrives and is received by the PEP. The PEP is the border component that receives access requests, redirects the request to the PDP, and enforces access execution after the evaluation results. The PIP gather attributes from internal or external sources to be evaluated by policies. Context Handlers CH translate raw attributes to contextual attributes, which PIP gathered [15]. ABAC is able to apply more fine-grained than RBAC and is also able to be used in complex systems in which RBAC cannot be applied [14]. The fact that ABAC uses a set of attributes more than just roles makes it better for vast scenarios with different complexity. Attributed-Based Access Control, especially using the XACML standard[16]. XACML can be used to define a set of logical rules which compose policies which verify whether the requestor can grant or deny access over the data [6].

2.3. Distributed Ledger Technology and Smart Contracts

Distributed Ledger Technology (DLT) is a digital storage technology for decentralising data held and updated for members in a network without a central authority. DLT is a group of decentralised technologies, and each has a way to store and share data. There are four types of DLT: DAG, Hashgraph, holochain and blockchain [17, 18].

Blockchain can be defined as a distributed ledger that stores data in a cryptographic way through chained blocks. The chain grows along more blocks as data is added[17]. Blockchain does not require a central authority to orchestrate the network; the nodes make decisions based on consensus, which are expressed in algorithms [18]. Blockchain uses well-known concepts and solutions from computer science like hash algorithms, public-key infrastructure, peer-to-peer networks, and Merkle tree. Those sets allow the blockchain to have the data inside the blocks in a way that allows all the data to be inserted on the leaves of the Merkle tree [19]. Additionally, this Merkle tree generates a hash based on the hash of each leaf. The hash of Merkle tree data is used to generate the hash of the actual block with other components. The hash of this block is generated using Merkle tree hash, a hash of the previous block, nonce and timestamp [20]. The blocks are stored entirely by network nodes in a way that if one or some nodes are unavailable, the data continues to be available because other nodes are online [21]. How many nodes store all chains of blocks makes the data available better than in a centralised system architecture. The structure of cryptographic chained blocks makes tampering with data computationally and mathematically infeasible.

The blockchain taxonomy can classify many different kinds of chains of blocks. Nevertheless, this work will be adopted as permissioned (or consortium blockchain) and permissionless. Permissionless are public blockchains in which anyone can be part of the network acting as a block validator, and all the data are open to anyone. On the other hand, permissioned blockchain allows only a permitted person or institution to be part of the validation of the network. Additionally, in the permissioned blockchain, the data is usually unavailable to the general public but to the partners of the network [22].

Smart contracts are self-executing scripts which run over a blockchain. Those scripts follow instructions to enforce the execution of operations on the decentralised ledger. Smart contracts are stored on the blockchain and wait to be triggered. Once a smart contract is called, it executes its functions and can store some data on the ledger [6]. Smart contracts can help develop decentralised applications over a blockchain, bringing reliability. In particular, it can happen by enforcing access and usage control over this environment, helping to have benefits such as transparency and auditability.

3. Related works

3.1. Centralised systems

Petroulakis *et al.* [23] propose a framework to exchange data in health data space. The work defines an architecture in which data spaces are decentralised and created tailored for fit purposes. Moreover, a centralised data space authority orchestrates and manages the data spaces and their sharing. The access control is made using keycloak. The drawbacks of this purpose are the centralised orchestration and centralised access control, which make auditability and sovereignty difficult.

Dam *et al.* [24] proposes a usage control mechanism for Mobility data spaces enforcing policies using Open Digital Rights Language (ODRL) to formalise policies. The interactions and data sharing are made through connectors linking data consumers and providers. The work has some enforcement mechanisms, such as prevention, detective, and continuous mechanisms. Prevent mechanisms act dynamically to allow or revoke access and also can delay access. Detective mechanisms act to prevent policy violations and detect misbehaviour. Continuous mechanisms ensure the obligations and conditions continue valid ongoing execution. The drawback of this process is the lack of auditability in data use and centralised take decisions.

Huang *et al.* [25] propose a usage control mechanism for Digital Rights Management (DRM) in a context in which a content provider outsources the storage and distribution of content for a cloud provider. Moreover, It implements usage control using homomorphic encryption and ABAC. The data is stored in an outsourced cloud, but the end users have access through their keys and attributes. The usage rights are encrypted with homomorphic encryption and stored in a cloud provider without the knowledge of the cloud of the content. The proposal addresses a challenge in the multimedia sector, but there is no auditability in the usage control.

Wang [11] proposes a trusted usage control framework for digital audio players that leverages Intel SGX to enforce secure playback policies and protect digital rights against piracy and reverse engineering. There are two components of the manage enclave: manages use, policy enforcement and authentication process. A player enclave still manages the audio files and decrypts them for use when the managed enclave authorises them. The proposal does not implement ABAC or XACML standards and has a centralised architecture.

Table 1
Comparison of Related Work

Work	Decentralised	Standard	Access Controlled by owner	Usage Controlled by owner	Auditability	Confidentiality
[6]	Yes	ABAC / XACML	Yes	✗	Yes	Yes
[11]	✗	✗	Yes	Yes	✗	Yes
[23]	Yes	✗	✗	✗	✗	Yes
[24]	✗	ODLR	Yes	Yes	✗	Yes
[25]	✗	✗	Yes	Partial	✗	Yes
[26]	Yes	✗	Yes	Yes	Yes	✗
[27]	Yes	ABAC	Yes	Partial	Partial	Yes
[28]	Yes		Yes	Yes	Partial	Yes
This Work	Yes	ABAC / XACML	Yes	Yes	Yes	Yes

3.2. Decentralised systems

Oliveira *et al.* [6] propose an access control mechanism to ensure safe access control and sovereignty of health care data. This paper uses blockchain, smart contracts, and XACML standards to ensure access control over the data. Access is allowed or denied according to the roles and context of the request based on attribute-based access control (ABAC). The drawback of this purpose is the lack of usage control.

Basile *et al.* [26] propose a usage control architecture to monitor compliance with usage control policies. The data share is made through SOLID, a tool for storing and sharing public and private data. The architecture is decentralised using blockchain, and each participant has its own data space orchestrated by a Trusted Environment Execution (TEE), which protects data. There is also a distExchange application that orchestrates the policies and data access and has the address of the source data. The TEE only allow access to data when a discharge permits. The drawback of the solution is confidentiality and privacy because the policies and resource locations are public, which means that a data space that has sensitive data can suffer attacks.

Denis *et al.* [27] propose a hybrid approach to address usage control on IoT data. There are central servers which store the data and a DAG DLT which stores the policies and makes the management of access usage control. The drawback of this approach is the lack of auditability in the design usage of DAG, which enables stakeholders in the network to be unaware of their partner's actions. It can give attackers space, leverage it, and exploit vulnerability in some nodes.

Siddiqui [28] proposes an access and usage control through blockchain using Hyperledger Fabric for Collaborative Mixed Reality (CMR). In mixed reality, the users share and consume virtual objects suck as 3D models. The system provides decentralised enforcement of usage policies and identifies malicious users through immutable activity logs. Although the proposal is decentralised and has auditability through the blockchain, there is no policy enforcement standard such as XACML or ABAC. The absence of an access control standard can make possible the enforcement of wrong definitions of policies and, as a result, leakage of data. Additionally, Hyperledger Fabric implements private channels, which, in some cases, can reduce transparency and auditability.

The works described above show approaches to address the access and usage control challenge in different fields. Some of them implement centralised solutions, and others decentralised with a DLT. Centralised solutions are not suitable because of the difficulty of ensuring auditability and data control by the owner; of course, it is possible, but not natively and not as safe as decentralised solutions with blockchain. On the other hand, decentralized solutions typically support access control, but often lack strong confidentiality protections of the data and standardized policy enforcement mechanisms. Standardized policy enforcement is essential to ensure that policies are correctly implemented, and the data is secure. Moreover, the gap in these works is the lack of a robust usage control mechanism that allows the data controller to track the data usage. The lack of better data usage control is a challenge we must address in research to bring trustworthiness to the whole system. It also helps to comply with GDPR requirements. Especially in decentralised environments such as data spaces or federated learning environments, It is essential to ensure an auditable and decentralized system for access and usage control, allowing stakeholders to maintain control over their data and track potential misuse. Finally, the data controller must ensure that data usage complies with defined access and usage policies, that data processors adhere to compliance rules, and that no data leakage or exposure occurs.

Table 1 compares related works. The table has decentralised columns, indicating whether the proposal uses decentralised architecture, especially with DLT. The column standard is about the access control and usage control type used, such as XACML or ABAC. Using that standard is important to ensure the security of the policies and if the policies are, in fact, enforcing what was designed. Moreover, the standards help to ensure the right enforcement and, as a result, data security. Access and usage controlled by data controller refers to whether the proposal design allows data controller to control access and usage of their data.

4. Research problem and motivation

In the data space environment, it is important to ensure that data controllers maintain control over their data and how third parties use it. It is necessary to give back control over access and usage of the data to the data controllers and ensure that data usage can be traceable to the data and that they can rule over usage. Access and usage control ruled by data controllers can bring more trustworthiness to the systems and guarantee laws obeyed, such as GDPR [2].

This research aims to ensure accountability, apply ethical and legal data sharing and usage, and raise trustworthiness among data-sharing participants in the data space. It is important in data-sharing environments to ensure compliance with the enforcing rules.

Research Questions

- RQ1: Is it possible with blockchain tracking, audit and detect when data is not being used as it was designed?
- RQ2: Is it possible to have an access and usage control mechanism that ensures traceability, accountability and controllability of the data and gives power to the data controller to revoke access when necessary?
- RQ3: Why do the related works not have a fully-tailored solution that implements access control and usage control, giving control to the data controller over the data hashing blockchain advantages, especially in a shared environment?
- RQ4: Does an access and usage control mechanism using blockchain help to track data leakage leveraging blockchain attributes (e.g.traceability);
- RQ5: Does a decentralised access and usage control mechanism with shared audibility bring trustworthiness among the stakeholders and end users (data subject)?
- RQ6: Are smart contracts enough to enforce and manage data usage control in real time, or is it computationally expensive to the blockchain?

Research Objectives

- RO1: To design a decentralised policy-based access control tailored for data sharing in data spaces using smart contracts to enforce policies;
- RO2: To enforce usage control policies tailored for decentralised data sharing in data spaces, focusing on controlling the data usage and tracking misuse of the data.
- RO3: Bring transparency and accountability in data usage through the distributed ledger;
- RO4: Evaluate the performance of the mechanism developed and if the metrics of scalability, throughput and time response are enough to use the mechanism in real scenarios such as industry applications. The evaluation can be done with some tools such as Hyperledger Caliper.
- RO5: Implement a security assessment using some tools such as Slither and Echidna.

In general, the mechanisms shown in related works do not have usage control as one of the main aspects or does not address the challenge (RQ3). Moreover, in a decentralised environment, when data processors want to share data between them, they need some guarantee that the data will be used as they agreed because of compliance with laws such as GDPR. In this direction, they must ensure that the data is being used according to the proposed proposed (RQ1). Additionally, data sharing in environments such as data spaces is important for tasks such as training ML models of federated learning (FL). In a data space with an FL environment, every stakeholder must ensure that the data are safe and can revoke access if necessary (RQ2) to track data leakage (RQ4). Moreover, because the environment is decentralised, it is necessary to have access and usage control traceable to the stakeholders and end users who can trust the whole environment (RQ5). However, the amount of data in a data space and the number of requests are considerable; it is necessary to have a self-executed mechanism such as smart contracts. However, we must analyse whether smart contracts can be resilient and scalable to this challenge (RQ6).

5. Proposed approach

European initiatives are trying to organise data use and sharing, bringing an environment safe to data sharing [10]. They have concerns about how the data is being used, profiling of the data subject and other ethical questions. Although the initiative seems to walk in the direction of preserving fundamental rights [29], protecting privacy without harm is the initiative of the companies. There is still a lack of an alternative to data access and usage control within this environment. Moreover, data spaces are an important initiative in the European Union for the development of Artificial intelligence (AI) and other technologies that need massive data. Usually, some sectors need cooperation between partners to improve results and predict failure or maintenance; the data space is essential, bringing a place where partners can collaboratively share data to improve the performance and results of the sectors. There are so many examples of this, and one of them is the aviation sector. An aviation company can preview the maintenance of an aircraft if it has a training model of its aircraft; the problem is that the sensor data are with the manufacturer. At the same time, the manufacturer does not build every component of the aircraft, and sometimes, they need the builder of a specific part to collaborate with data to create a new aircraft model with cost reduction. Another example can be the health care system, how to ensure data sharing between hospitals of a patient without harming the patient and the hospital provider, or between a hospital and a third party that needs a specific part of data but does not have access to the whole data. How do we ensure that the third party will process the data exactly as previously agreed?

There are other problems. A few companies, such as big techs, control an enormous mass of the data and do what they want without transparency and accountability, even with novelty regulations such as GDPR [30]. However, the regulations are important; enforcing the principles and fundamental rights in the systems is necessary. Another aspect is the power of big tech companies worldwide to control trends and political discussions without users' consent. Additionally, there is a security problem; if an entity controls massive data, if leakage happens, that big mass of data is compromised [31]. Thus, in this case, decentralisation is important to the privacy and security of the data; even if a data controller is compromised, only a part of the whole data in a shared environment, such as data space, is not compromised.

The proposed research aims to design and implement a decentralised access and usage control mechanism in data spaces, leveraging Distributed Ledger Technology (DLT) and smart contracts for enforcing policies. The proposal design integrates Attribute-Based Access Control (ABAC) policy definitions with XACML standards and enforces them through smart contracts deployed on a blockchain.

This work aims to address some of the challenges that the mechanisms in section 3 did not address. The proposed mechanisms do not present a robust and tailored access and usage control mechanism for shared environments, especially when this environment is decentralised and needs traceability, trustworthiness and compliance. There is a gap in ensuring that the data is being processed and used according to the policies defined, and the data controller can revoke access and use of the data in case of misuse. Additionally, it is important to ensure compliance by following laws such as GDPR. Data sharing is essential for tasks requiring massive data, such as federated learning. However, enforcing policies that ensure the data subject's privacy (end-user) and correct data access and usage is important to bringing trustworthiness to the data-sharing environment.

This paper proposes a mechanism of access and usage control of data in a shared environment (RO2), ensuring control of the data by the data processor and giving them the power to track improper access and misuse (RQ1). We are trying to do it through blockchain, implementing policy enforcement through smart contracts (RO1) to allow or deny access to the data and also controlling the data usage by leveraging the accountability (RO3) feature of the blockchain to address the data usage in shared environments such as data spaces (RQ2).

This research aims to be flexible and can integrate the proposal into other research challenges requiring access and usage control, such as federated learning for aviation. In aviation, there is the need to share training models with partners that have data to use their data to improve the training models. In this scenario, there are some challenges and rewards for trustworthiness. This does not mean that a reward mechanism exists in every access and usage control for data space usage. Moreover, data access and usage control can be used in different scenarios within data spaces; one of the examples is an example of federated learning.

To describe the architecture of the proposal, we are using the example of a federated learning training model that uses data space, which needs access and usage control for the data. The purpose is to present the proposal's feasibility within a scenario for the best understanding. Figure 1 presents the architecture in which a model is trained in a federated learning approach. The data controller sets the policies for the model's usage and access control and makes the model available for training. Data processors can improve the training model, request training rewards according to contribution, and receive compensation. Below, we have the mechanism described in steps and integrated with the reward mechanism. An important aspect is that in this approach, the reward mechanism is a smart contract that also works according to the contribution of each data processor.

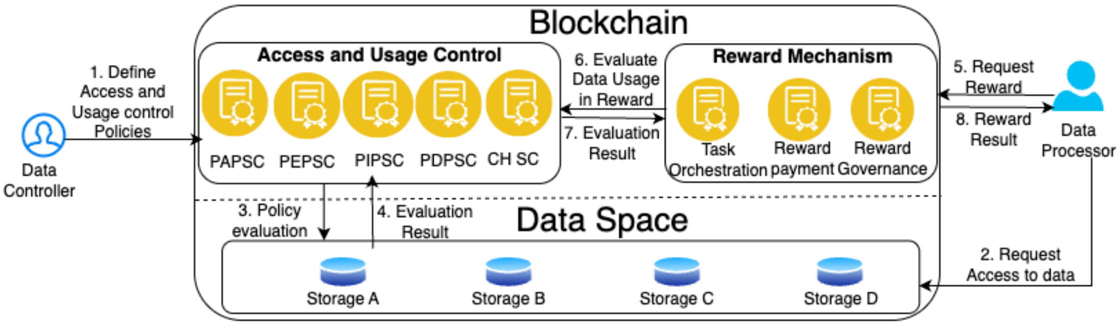

Figure 1: Architecture proposal

Figure 1 shows the architecture and the sequence of actions step by step and is described below.

Step 1: The Data Controller defines and deploys access and usage control policies onto the blockchain. These policies are enforced through a suite of smart contracts.

Step 2: A Data Processor requests access to data stored in the Data Space. The request is sent to PEPSC.

Step 3: The request triggers a policy evaluation process to PEPSC and sends the request to PDPSC, where the access and usage control components are called. The PDPSC component has smart contract addresses that store and evaluate policies together, such as PAPSC and PIPSC. The contract Policy Administration Point (PAP) stores and manages access and usage control defined by the data controller. PAPSC recover the policy and delivers it to the PDPSC according to the resource requested and the user.

Step 4: PIPSC: After the PDPSC request, interact with the context handler (CHSC), which is responsible for gathering off-chain data that cannot be stored in the ledger. The PIPSC gathers data on the context attributes and returns the context policy with the context attributes to the PDPSC.

Step 5: PDPSC gathers the policies and information provided by PAPSC and PIPSC and evaluates them to produce a result.

Step 5: The policy evaluation result is returned, which determines whether the consumer can proceed to the PEPSC.

Step 6: The PEPSC emit the result to the data processor to allow or deny access and usage of the data.

Step 7: If the data processor performs an action that qualifies for a reward, the rewarding process is triggered when it can be applied following the access control flow.

Step 8: The Reward Mechanism receives a usage evaluation request and interacts with the access and usage control logic to assess if the consumer met the expected conditions.

Step 9: The evaluation result is sent back to the Reward Mechanism.

Step 10: If conditions are satisfied, a reward is issued and returned to the data consumer.

This work does not detail reward mechanism aspects such as the reward token, access, or other kind of reward because it is the subject of other research. Those steps can be changed according to the research context and evolution. Thus, the steps and the diagram are starting points to address the research questions.

6. Expected Contributions and Research plan

A novel integration model that combines ABAC and usage control in a decentralized architecture using smart contracts. An extension of the XACML policy to enable on-chain enforcement and usage policies representation auto-executed. A conceptual architecture for decentralized enforcement of fine-grained access and usage control tailored for data spaces, addressing trust and accountability challenges. We also have a plan to develop a prototype implementing smart contract–based policy enforcement mechanism, deployable on permissioned DLT platform Hyperledger Besu. A validation plan using data sharing scenarios to evaluate policy enforcement in terms of performance and scalability (RO4). A security evaluation to assess the security of the smart contracts (RO5).

The proposed research is structured over 4 years. The ongoing activities focus on conducting a comprehensive literature review and defining the architecture for a decentralized access and usage control mechanism tailored for data spaces. The methodology to develop the research follows a design science approach. The next phases will be carried out according to the following plan:

- **Architecture Design and Policy Modeling**
  Finalize the definition of the overall system architecture, including integration between components and innovative contract policies within a permissioned DLT environment. Develop the initial policy models for both access and usage control.
- **Implementation and Prototype Development**
  Implement the proposed smart contract–based control mechanism using Hyperledger Besu (RO1), (RO2). Develop and test the policy translation and enforcement logic, ensuring interoperability with standard policy languages. Define and integrate the proposal with real use cases like the aviation sector (RO3).

- **Evaluation and Validation**
  Define test scenarios. Evaluate performance using Hyperledger Caliper within defined test scenarios, measuring latency, throughput and resource consumption (Ro4). Security evaluations will also be carried out through static code analysis with Slither and *fuzz testing* of smart contracts using Echidna and Foundry to assess vulnerabilities (RO5). Fix vulnerabilities found in smart contracts.

7. Conclusion

Access and usage control are challenges in these years, mainly because of the advent of big data and AI. In this way, concerns about data usage and access have been raised. Although some legal actions were taken, other aspects and approaches must be taken to address the challenge. One of the actions is to bring back to data controllers the sovereignty of their data. This paper is a research proposal for PhD to address this challenge by using a decentralised approach with access and usage control patterns such as ABAC and XACML to design, implement and enforce policies developing and using smart contracts (RO1), (RO2), (RO3). The plan for this research considers theoretical and practical challenges that we must address to reach the goals based on the motivations defined. The expectation is that after the development of the proposal and its deployment in a blockchain, some experimentation, such as performance and security analysis, will be carried out (RO4), (RO5). We plan to improve the architecture and, right after, code the smart contracts and deploy them in a permissioned blockchain such as Hyperledger Besu. After the implementation, we must assess the smart contract's performance and security.

The performance analysis evaluates and analyses the behaviour of the proposal in a test environment and, if possible, a real scenario, such as a company environment or research project. Usually, performance assessment evaluates throughput, latency and resource consumption as performance assessment metrics. However, this research will define the metrics in other opportunities. On the other hand, it is possible to use a hyperledger caliper for performance assessment, which is a tool used for it. Nevertheless, it is a decision that will made in the future.

The security assessment helps analyse the security of decentralised solutions against attacks over the network. The reason for security analysis is that if the smart contracts were compromised, the data could be leaked or have unauthorised access. Security analysis can be done using tools such as Slither for static analysis and Echidna for fuzzy and invariant tests. The reason for doing a security assessment is that the access and usage control mechanism should be safe; otherwise, the data usage could be compromised.

Acknowledgments

This study was partially funded by the European Union's Horizon Program under grant agreement no. 101093164 (ExtremeXP: EXPerimentation driven and user eXPerience oriented analytics for eXtremely Precise outcomes and decisions).

Declaration on Generative AI

During the preparation of this work, the author(s) used ChatGPT-4, Grammarly in order to: Grammar and spelling check, Paraphrase and reword. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the publication's content.

References

[1] European Parliament and Council, Regulation (eu) 2024/1689 of the european parliament and of the council of 13 june 2024 on harmonised rules on fair access to and use of data (data act), 2024. URL: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32024R1689, accessed: 2025-04-25.

[2] European Parliament and Council, Regulation (eu) 2016/679 of the european parliament and of the council of 27 april 2016 (general data protection regulation), 2016. URL: https://eur-lex.europa.eu/eli/reg/2016/679/oj/eng, accessed: 2025-04-25.

[3] European Union, Regulation (eu) 2024/1689 of the european parliament and of the council, 2024. URL: https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=OJ:L_202401689, accessed: 2025-05-20.

[4] European Commission, What is a data controller or a data processor?, https://ec.europa.eu/info/law/law-topic/data-protection/reform/rules-business-and-organisations/obligations/controller-processor/what-data-controller-or-data-processor, 2016. Accessed: 10/04/2025.

[5] European Commission, Joint controllers, https://gdpr-info.eu/art-26-gdpr/, 2016. Accessed: 10/04/2025.

[6] M. T. De Oliveira, L. H. A. Reis, Y. Verginadis, D. M. F. Mattos, S. D. Olabarriaga, Smartaccess: Attribute-based access control system for medical records based on smart contracts, IEEE Access 10 (2022) 117836–117854.

[7] International Data Spaces Association, International data spaces association (idsa), 2020. URL: https://internationaldataspaces.org/, accessed: 2025-04-24.

[8] European Union, What is a gdpr data processing agreement?, https://gdpr.eu/what-is-data-processing-agreement/, 2016. Accessed: 10/04/2025.

[9] European Commission, European strategy for data, 2020. URL: https://digital-strategy.ec.europa.eu/en/policies/strategy-data, accessed: 2025-04-23.

[10] European Commission, Common european data spaces, 2020. URL: https://digital-strategy.ec.europa.eu/en/policies/data-spaces, accessed: 2025-04-23.

[11] J. Wang, Research on manufacturing data space storage based on data space, in: 2022 IEEE International Conference on Advances in Electrical Engineering and Computer Applications (AEECA), IEEE, 2022, pp. 106–110.

[12] European Commission, Smart middleware for trusted data spaces (simpl), 2024. URL: https://digital-strategy.ec.europa.eu/en/policies/simpl#1712822729753-0, accessed: 2025-04-25.

[13] T. Dam, A. Krimbacher, S. Neumaier, Policy patterns for usage control in data spaces, arXiv preprint arXiv:2309.11289 (2023). URL: https://arxiv.org/abs/2309.11289.

[14] P. Hlushchenko, V. Dudykevych, Exploratory survey of access control paradigms and policy management engines, in: Proceedings of the Seventh International Workshop on Computer Modeling and Intelligent Systems (CMIS-2024), volume 3702 of *CEUR Workshop Proceedings*, 2024, pp. 192–202. URL: https://ceur-ws.org/Vol-3702/paper22.pdf.

[15] M. T. de Oliveira, Y. Verginadis, L. H. Reis, E. Psarra, I. Patiniotakis, S. D. Olabarriaga, Ac-abac: Attribute-based access control for electronic medical records during acute care, Expert Systems with Applications 213 (2023) 119271.

[16] OASIS, eXtensible Access Control Markup Language (XACML) Version 3.0, http://docs.oasis-open.org/xacml/3.0/xacml-3.0-core-spec-os-en.html, ???? Accessed: 30/06/2025.

[17] E. Ribeiro, J. Soares, et al., A method for quality evaluation of supervision software using fuzzy concepts and the international standard iso/iec 25000, Journal of Control, Automation and Electrical Systems 28 (2017) 389–404.

[18] A. Panwar, V. Bhatnagar, Distributed ledger technology (dlt): The beginning of a technological revolution for blockchain, in: 2020 2nd International Conference on Data, Engineering and Applications (IDEA), IEEE, 2020, pp. 1–5. doi:10.1109/IDEA49133.2020.9170672.

[19] M. Belotti, N. Božić, G. Pujolle, S. Secci, A vademecum on blockchain technologies: When, which, and how, IEEE Communications Surveys & Tutorials 21 (2019) 3796–3838. doi:10.1109/COMST.

2019.2928178.

[20] A. Alketbi, Q. Nasir, M. A. Talib, Blockchain for government services—use cases, security benefits and challenges, in: 2018 15th Learning and Technology Conference (L&T), IEEE, 2018, pp. 112–119. doi:10.1109/LT.2018.8368491.

[21] K. Yue, Y. Zhang, Y. Chen, Y. Li, L. Zhao, C. Rong, L. Chen, A survey of decentralizing applications via blockchain: The 5g and beyond perspective, IEEE Communications Surveys & Tutorials 23 (2021) 2191–2217. doi:10.1109/COMST.2021.3105233.

[22] G. Falazi, M. Hahn, U. Breitenbücher, F. Leymann, V. Yussupov, Process-based composition of permissioned and permissionless blockchain smart contracts, in: 2019 IEEE 23rd International Enterprise Distributed Object Computing Conference (EDOC), IEEE, 2019, pp. 77–87. doi:10.1109/EDOC.2019.00017.

[23] N. Petroulakis, P. Zervoudakis, G. Nomikos, A. Kornilakis, P. Chatziadam, D. Laskaratos, V. Maria-Eleftheria, Z. Eleni, V. Theodorou, Towards the development of a network provisioning platform for data exchange in the health data space, in: 2024 IEEE Conference on Standards for Communications and Networking (CSCN), IEEE, 2024, pp. 147–153.

[24] T. Dam, A. Krimbacher, S. Neumaier, Policy patterns for usage control in data spaces, arXiv preprint arXiv:2309.11289 (2023).

[25] Q. Huang, Z. Ma, Y. Yang, X. Niu, J. Fu, Attribute based drm scheme with dynamic usage control in cloud computing, China Communications 11 (2014) 50–63.

[26] D. Basile, C. Di Ciccio, V. Goretti, S. Kirrane, A blockchain-driven architecture for usage control in solid, in: 2023 IEEE 43rd International Conference on Distributed Computing Systems Workshops (ICDCSW), IEEE, 2023, pp. 19–24.

[27] N. Denis, M. Laurent, S. Chabridon, Integrating usage control into distributed ledger technology for internet of things privacy, IEEE Internet of Things Journal 10 (2023) 20120–20133.

[28] M. S. Siddiqui, Activity provenance for rights and usage control in collaborative mr using blockchain, in: 2023 20th Learning and Technology Conference (LT), IEEE, 2023, pp. 26–31. URL: https://ieeexplore.ieee.org/document/10092326. doi:10.1109/LT58159.2023.10092326, accessed via TU Delft Library.

[29] European Union, Gdpr art. 6 lawfulness of processing, https://gdpr-info.eu/art-6-gdpr/, 2016. Accessed: 10/04/2025.

[30] K. Paul, How amazon tracked my last two years of reading, 2020. URL: https://www.theguardian.com/technology/2020/feb/03/amazon-kindle-data-reading-tracking-privacy.

[31] R. Josphineleela, S. Kaliappan, L. Natrayan, A. Garg, Big data security through privacy–preserving data mining (ppdm): A decentralization approach, in: 2023 second international conference on electronics and renewable systems (icears), IEEE, 2023, pp. 718–721.