

Using Blockchain to Secure Passenger Name Records and Effectively Protect Citizen Rights

Michal Koutenský
Vladimír Veselý
Martin Perešíni
Daniel Dolejška
Jan Pluskal

*Faculty of Information Technology
Brno University of Technology
Brno, Czechia*

Abstract—Passenger Name Records (so called PNR) collected by airlines and processed by Passenger Information Units (dedicated police task force), are critical data in the prevention and successful investigation of terrorism, drug smuggling, human trafficking and other serious crimes. This work addresses the challenges of building a blockchain-based, tamper-proof system for PNR data exchange between different Passenger Information Units. Utilizing Hyperledger Fabric, we propose a decentralized architecture that ensures data confidentiality, traceability, and non-repudiation in order to maintain the chain of custody. Our solution includes a private data collection mechanism and an off-chain storage system to manage sensitive information securely. The resulting design and implementation are incorporated into the TENACITY project platform and further integrated with its other modules to use travel intelligence data in the fight against crime. The blockchain subsystem facilitates a secure, auditable exchange of PNR data, with the key benefit being compliance with European Union regulations on the protection and retention of personal data. This paper contributes to the field by demonstrating the practical application of blockchain technology in a sensitive data exchange scenario, offering insights into system design, implementation challenges, and potential future enhancements.

Index Terms—blockchain, hyperledger fabric, smart contract, passenger name record, advanced passenger information, travel intelligence

I. INTRODUCTION

Since the year 2016, the European Union (EU) has adopted several laws (the most relevant would be [1], [2]) that harmonize the approach to collecting and processing data about people flying in Europe. The main goal of this policy is to effectively prevent, detect, investigate and prosecute serious criminal activities (such as terrorism, drug smuggling or human trafficking). Air carriers operating within the EU are required to collect Passenger Name Record (PNR) data and Advanced Passenger Information (API), both containing personal information provided by passengers as part of the booking and check-in process. PNR consists of the passenger's name, frequent flyer information, reservation and travel dates, itinerary (i.e., a list of airports visited), seat number(s),

baggage information (e.g., how many bags, if any, and their weight), contact information (e.g., phone number, e-mail address), payment method (e.g., direct purchase, travel agent payment), billing information and others [3]. The API contains more reliable data obtained from official travel documents (e.g., passport, citizenship card), such as full name, date of birth, nationality, and possibly other personally identifiable information (PII). PNR and API are analyzed mainly for these reasons: a) identification of criminals and their associates; b) detection of suspicious travel patterns; c) passenger due diligence against defined risks; and d) development of heuristics and risk criteria.

Legislation mandates the establishment of a dedicated Passenger Information Unit (PIU) in each EU member state. The PIU is responsible for collecting, processing and disseminating the data to other PIUs, National Competent Authorities (NCA), Europol or, to a limited extent, with other involved law enforcement agencies (LEA) outside EU. This cross-unit exchange of PNR data is paramount to counter criminal activities. However, PNR data—by its nature—is sensitive personal information and needs to be handled accordingly to preserve the right to privacy of citizens traveling through EU. As such, PNR data must be depersonalized after a few months and deleted after a few years so that the passenger cannot be immediately identified [4]. In addition, passengers should be clearly informed that their PNR data are being gathered and of their rights related to this collection process.

The Travel Intelligence Against Crime and Terrorism (TENACITY) project¹ aims to develop a modern set of tools for risk analysis and crime prevention. These tools are all meant to work with the local PNR data available to each PIU. In certain cases, it might be useful to request additional PNR data (e.g., followup flights from destination airport) to better assess the situation. Within the TENACITY project, this data exchange is facilitated by a blockchain subsystem, integrated with the rest of the toolbox.

The main contribution of this paper is to describe the design

This project has received funding from the European Union's Horizon Europe research and innovation program under the topic HORIZON-CL3-2021-FCT-01-01 with grant agreement No. 101074048

¹For more information about the project, its goals and deliverables, please, visit <https://tenacity-project.eu/>

and implementation of a blockchain subsystem that enhances PNR data exchange between PIUs with inherent distributed ledger properties such as decentralization, auditability, and built-in anti-tamper protection. The goal of this blockchain subsystem is to provide a secure communication channel for requesting and providing PNR data between two PIUs. Each request (and the associated response) should be permanently recorded so that any actions taken (within the PNR exchange workflow) are non-repudiable. However, the requested PNR data itself must never be made available to any other party than the requesting PIU. Furthermore, all PNR data must be depersonalized and deleted after a certain time period.

This article is organized as follows. Section II outlines related work and the state of the art of the technologies used. Section III describes the main use case for the blockchain subsystem, which is secure exchange of PNR data, in more detail. Section IV informs about the design of the blockchain subsystem and its implementation and integration with other TENACITY components. Section V discusses the achieved state and future work.

II. RELATED WORK

The adoption of blockchain technologies has gained significant traction across numerous sectors due to its promise of enhanced security, transparency, and efficiency. Although there are concerns whether blockchain technologies can deliver on their promises of truly decentralized, self-governing organizations [5], they can be utilized as building blocks for distributed applications that require a tamper-proof way of record keeping. The EU has adopted blockchain for the public sector, in the form of European Services Blockchain Infrastructure [6]. As blockchain technologies have the potential to be wildly disruptive, the EU has also published a set of ethical guidelines [7] for developers to consider when designing new systems.

Various blockchain technologies have emerged over the years, differing in features and targeting diverse use cases. The problem of selecting the right blockchain technology has been studied by several authors, with a number of methodologies proposed [8]–[10]. Applying the methodology in [8], our use case requires a *full-permissioned blockchain*. This section therefore reviews pertinent application developments focusing on permissioned blockchain technologies like Hyperledger Fabric [11], Hyperledger Besu [12], Quorum [13], and Corda [14]. We also explore the concept of private data collection within these frameworks and their possible use cases. In addition, we provide relevant information about the existing PNR sharing system.

A. Hyperledger Fabric

Hyperledger Fabric [11] is a prominent enterprise-grade blockchain framework that supports a modular architecture, allowing for a diverse range of industry-specific applications while ensuring high privacy and confidentiality. Noted for its permissioned nature, Fabric utilizes identity services to authenticate and authorize participants, unlike permissionless blockchains such as Ethereum or Solana. One of the defining

features of Hyperledger Fabric is its use of private data collections, which allow data to be segregated between subsets of network participants. This ensures that sensitive information is not exposed to all nodes in the network, a crucial feature for applications requiring strict data privacy and regulatory compliance.

Private Data Collections

Private data collections [15] in Hyperledger Fabric provide a sophisticated method for handling sensitive information. When a private data collection is created, it ensures that only the nodes of participating organizations can store and see the data. This is particularly beneficial for healthcare, finance, and supply chain management industries, where sensitive data must be protected from unauthorized access.

Private data collections enable organizations to maintain data privacy and comply with regulatory requirements by isolating confidential information. They also allow for fine-grained control over who can see and interact with certain pieces of data, facilitating secure and compliant business processes. In addition, the endorsement policies for private data collections can be customized, allowing organizations to specify which peers must endorse transactions before they are committed to the ledger.

B. Alternatives

Hyperledger Besu: Hyperledger Besu [12], developed by ConsenSys and now part of the Hyperledger project, is an Ethereum client designed for both public and private permissioned networks. Supporting various consensus mechanisms such as Proof of Work (PoW), Proof of Authority (PoA), Quorum and Istanbul Byzantine Fault Tolerance (QBFT, IBFT), Besu is a versatile platform for blockchain applications. Although Besu was built from scratch and is not a fork of go-ethereum (official client), it is fully compatible with the Ethereum Virtual Machine (EVM). It offers robust privacy features, including the ability to perform private transactions using Tessera [16] nodes, which are crucial for businesses that wish to comply with data protection regulations while leveraging the public Ethereum ecosystem, in a similar manner as Private Data Collections in Hyperledger Fabric, but using a different technique. Tessera operates by encrypting transaction payloads and managing the distribution of these encrypted payloads to the relevant parties. When a private transaction is created, the transaction payload (the actual data and instructions) is not broadcast to the entire network; instead, it is encrypted and sent only to the particular relay nodes involved in the transaction. Tessera ensures that only the authorized participants receive the encrypted transaction payload.

Corda: Corda [14] is an open-source blockchain platform designed for businesses to transact directly and in strict privacy using smart contracts. Unlike typical blockchain platforms, Corda does not employ a global shared ledger. Instead, it ensures that data is shared only between parties with a legitimate need to know, making it highly suitable for complex transactions and workflows, particularly in the financial services

industry. While Corda excels in providing a highly private and direct transaction model suitable for financial services, we chose Hyperledger Fabric with private data collections for several key reasons such as Hyperledger Fabric's modular architecture, broader industry support, endorsement policies and lastly wide community and ecosystem of working projects in Hyperledger framework.

Quorum: Quorum [13] is an enterprise-focused version of Ethereum designed for applications requiring high speed and high throughput processing of private transactions within a permissioned group of known participants. Quorum addresses specific challenges to blockchain technology adoption within the financial industry and beyond.

Each of these platforms provides robust frameworks for businesses looking to leverage blockchain technology to address specific needs, particularly in terms of privacy, security, and efficiency. Each brings unique strengths to various applications, underscoring the flexibility and potential of blockchain as a transformative tool for digital transactions, and each has its unique deployment and niches.

Privacy-Preserving smart contract Platforms (PPPs)

In addition to the existing blockchain frameworks mentioned above, PPPs offer a unique approach specifically tailored for data privacy. These platforms leverage advanced cryptographic techniques, often employing a combination of fully homomorphic encryption (FHE), secure multi-party computation (SMPC), or zero-knowledge proofs (ZKPs), to shield data and guarantee privacy. They often operate in conjunction with trusted execution environments (TEEs) for enhanced security. The following are some notable privacy-preserving platforms that can be applied to the use case of data exchange:

Secret Network: [17] This platform utilizes TEE to enable smart contracts that keep all data (inputs, outputs, and states) encrypted. This ensures sensitive information within smart contracts remains confidential, making it ideal for applications in healthcare and finance where data privacy is paramount.

Oasis Network: [18] Oasis Network prioritizes data privacy and confidentiality through its secure computation (powered by TEE) and decentralized infrastructure. The platform facilitates confidential smart contracts that maintain data privacy while enabling computations to be verified on the blockchain.

Phala Network: [19] Similar to the above, Phala Network offers privacy-preserving cloud computing services integrated with blockchain technology. It utilizes TEEs to establish a secure execution environment for smart contracts, making it well-suited for applications like confidential data analysis.

Beyond data exchange, PPPs have other potential use cases such as data sharing platform [20] and electronic voting (e-voting) [21]. In e-voting, PPPs can enhance accessibility by leveraging encrypted proofs that guarantee ballot integrity for voters. However, challenges remain. Verifying voter identity and scaling PPPs to handle nationwide elections pose significant downsides, as these platforms rely heavily on computationally intensive cryptographic constructs, making them less efficient for large-scale applications.

C. Existing Use Cases

As previously mentioned, the application of blockchain technology, particularly Hyperledger Fabric, has been explored across various sectors (besides finance and cryptocurrencies) to build decentralized systems requiring data integrity, traceability, data provenance and access control.

Healthcare Data Security and Identity Management: Providing proper healthcare services requires handling of sensitive patient data and the exchange of this data between various stakeholders, such as the healthcare providers themselves, insurance companies, patient's relatives, etc. Use of blockchain is being explored to address the challenges associated with the reliability, security and data ownership of such systems. Beinke et al. [22] gathered stakeholder requirements and proposed a possible architecture for a blockchain-based Electronic Health Record system. Antwi et al. [23] implemented a solution based on Hyperledger Fabric and evaluated it on test case scenarios, reaching a positive conclusion about the applicability of blockchain.

Data Exchange and Transparent Logging: Blockchain technology is also applicable to data exchange services. Hoye et al. [24] showcase the use of Hyperledger Fabric for secure data sharing in a cross-organizational context. Their approach leverages Fabric's logging attribute to prevent disputes during data exchanges and facilitate the framework for the creation and termination of short-term collaborations (ad-hoc sharepoints). For transparent logging, Schaefer and Edman [25] propose a hybrid blockchain architecture. This architecture combines Hyperledger Fabric with a public permissionless blockchain to capture and log data for each customer. This approach ensures data provenance and transparency in personal data handling. Furthermore, blockchain frameworks are increasingly used for Internet of Things (IoT) data sharing. For instance, IOTA [26] is a permissionless blockchain specifically designed for IoT applications. In the context of permissioned systems, Chen et al. [27] explore the use of Hyperledger Fabric for secure and privacy-preserving data sharing within Industrial Internet of Things (IIOT) enterprises, reaching a solution similar to ours. They utilize the channel mechanism to enable enterprises to share data while keeping sensitive data private and communication on different channels isolated. At the same time, all transaction data are timestamped and permanently stored in the blockchain ledger (ensuring non-repudiation).

D. Secure Information Exchange Network Application

Collecting PNR from air carriers is being done on a national level by various set of interfaces for automated data exchange. There are handbooks specifying the process of enrolling and subsequent transfer of information between air carrier and its corresponding PIU (e.g., guide for the Czech Republic [28], which supports data exchange via dedicated web portal, RESTful web interface, and couple of proprietary message queuing systems). PNR data can be stored in various formats including PNRGOV EDIFACT [29] or PNRGOV XML [30], where the exchange is administered via a sequence of request-response-confirmation messages with standardized content.

Since its launch in 2009, the Secure Information Exchange Network Application (SIENA) has been one of Europol’s core applications for LEAs coordination. SIENA allows swift exchange of crime-related information between EU LEAs, Frontex, Eurojust, OLAF, Interpol and cooperating countries outside EU (such as USA, Switzerland, Canada)—in total 51 countries and 14 international organizations. SIENA is an information system supporting data exchange (guaranteeing integrity and confidentiality) between users (associated with various roles and permissions) [31]. It is used as the primary platform for PNR data exchange between PIUs (i.e., one PIU requesting PNR from another PIU).

From the point of view of EU citizens, SIENA platform has following two properties:

- Based on publicly available information, SIENA closely follows EU regulations and is compliant with legal requirements for data protection and confidentiality. For obvious reasons (i.e., sensitive nature of information for day-to-day operations of various LEAs), general public does not have an access to SIENA. As such, citizens do not have direct evidence about proper handling of their personally identifiable information by SIENA parties according to the PNR data retention policies of GDPR compliance.
- SIENA is a centralized solution operated by a single party (i.e., Europol). Hence, SIENA availability could be compromised or data integrity/confidentiality inherently questioned.

SIENA is used as a de facto industry standard for the exchange of structured and unstructured data between LEAs (including PIUs). SIENA provides only indirect evidence (based on which organizations are involved in its operation) of its compliance (when it comes to handling citizens’ PII), as details of its implementation and technology stack are proprietary. It is therefore useful to consider an alternative to SIENA that: a) is cryptographically secure by design; b) supports processes of PIUs; c) follows the EU regulation on the retention of travel-related PII; and d) has publicly available source code. This is precisely the main driver behind the design and implementation of our blockchain subsystem. Due to its proprietary nature, it is difficult for us to conduct a comparative analysis of SIENA with our work. Nevertheless, our intention is not to replace SIENA, but to provide an alternative exchange channel with interesting built-in features from the perspective of both PIUs as well as citizens.

III. PNR EXCHANGE USE CASE

The TENACITY project aims to build a holistic, comprehensive toolbox for crime prevention and risk management. The tools and their connections can be seen in Fig 1.

The central point is the Risk Management Tool (RMT). It ingests PNR and API data from external data sources (such as PIU database or air carrier) and provides a central data repository for the local TENACITY instance. Other tools, as well as RMT itself, then use this repository for their specific use cases.

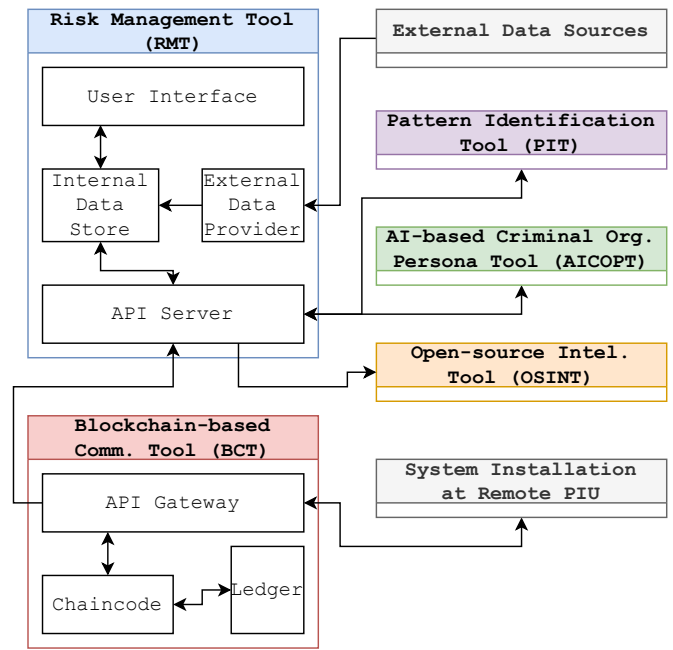


Fig. 1. Architecture of the TENACITY toolbox and its subsystems.

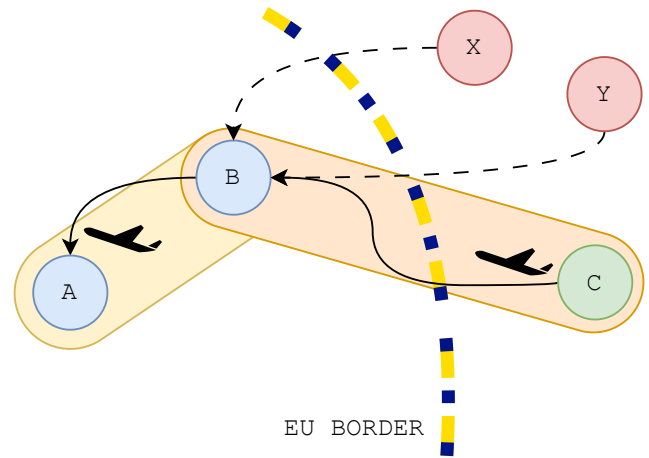


Fig. 2. Example risk assessment scenario showing the motivation behind PNR exchange use case.

It should be noted that this RMT data ingestion is only used for data which would be nonetheless available to the particular PIU unit. In certain cases, it might be helpful to know information that the PIU itself does not have—but another PIU does.

Consider a passenger flying from non-EU airport *C* to EU airport *A*, with transit stop at EU airport *B*, as shown on Fig. 2. The PIU of country *A* will only have information about the flight from *B* to *A* (yellow). In the case an officer determines that there is reasonable suspicion that a passenger matches the suspect profile in one of their investigations, they might require further information to validate this suspicion. In this scenario, our investigation suspect is known to fly to EU from airports *X* and *Y* (shown in red). The origin airport of flight

to *B* (green) is not available to PIU *A*, but it is known to PIU *B*, as they were the transit point. An officer of PIU *A* would therefore like to request this information to help in their investigation.

From a technical standpoint, there are several requirements for this kind of data exchange that need to be upheld. Firstly, as we are dealing with sensitive personal data, it is paramount that the transfer is fully confidential and no unauthorized party can gain access to the data being exchanged. Secondly, actions taken during the exchange process need to be traceable and non-repudiable. This helps prevent malicious actors from abusing the system, as well as alleviate resolving disputes in case of human error. Lastly, the system should be decentralized, so that no member party has a privileged position with regards to the operation of the system and the constraints it enforces.

IV. DESIGN AND IMPLEMENTATION

The requirements identified in section III align nicely with properties provided by blockchain technologies. Blockchain networks are decentralized, on-chain information is traceable and non-repudiable; what is missing is data confidentiality and access control.

A. Technology choices

In blockchain networks that support smart contracts, these properties can be implemented on top of the base blockchain. An off-chain storage stores the actual data in an encrypted form, and the blockchain stores the necessary information (i.e., hash of the contents) to be able to verify the the authenticity of the data. Access control is handled in the smart contract or the off-chain storage. We have previously used a similar approach (private Ethereum network with IPFS for off-chain storage) as part of the BLENDED project [32]. There are also ready-made solutions, such as Quorum or Hyperledger Besu, which can provide private transactions on top of Ethereum by utilizing Tesseract nodes.

While it is great that there are options within the Ethereum ecosystem which attempt to provide such features, extending a system in such manner comes with downsides. These supporting systems (IPFS, Tesseract) need to be deployed, configured, managed, and in the case of IPFS, manually integrated. Working with an existing system results in technical limitations, such as Quorum private transactions being unable to write to public contracts. As Ethereum was designed to be a public network, it is sometimes necessary to work around features which might not be necessary in private, permissioned networks, e.g., transaction fees (gas). Both Besu and Quorum allow creation of gas-free networks, but Ethereum (geth) does not.

As we are fortunate enough to be starting from a clean slate, with no compatibility requirements, building a closed, single-purpose network, we have decided to look for technologies which would better fit our needs. The Hyperledger umbrella covers a variety of blockchain related projects, targeting different use cases. One of these is Hyperledger Fabric, a permissioned blockchain for industry.

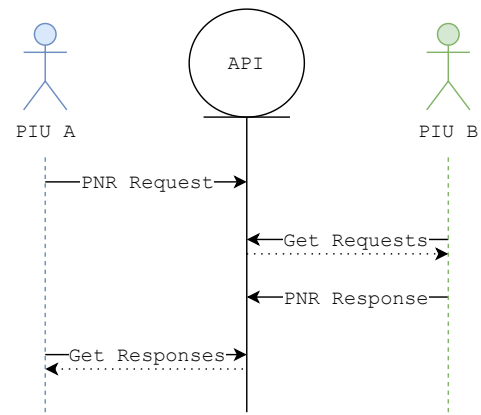


Fig. 3. Conceptual diagram of the PNR exchange use case.

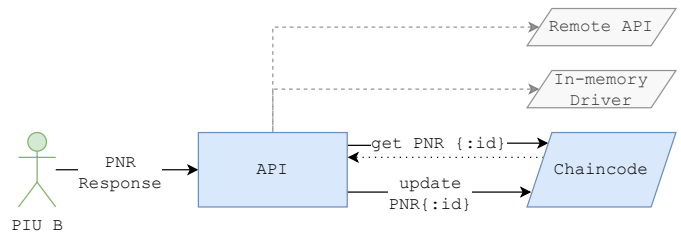


Fig. 4. Architecture of the blockchain subsystem, showing possible data repository backends.

As a permissioned network, Fabric contains an identity layer based on TLS certificates that allows modeling organizational hierarchies. Every organization within a single network can participate in several channels, each with its own configuration, smart contracts and ledger. There are no transaction fees, with dedicated orderer nodes ensuring network consensus without Proof-of-Work/Proof-of-Stake, utilizing consensus algorithms such as Raft. Last but not least, Fabric has a built-in concept of private data that is only shared with a certain set of organizations.

All of these features are something that we either require, or give us flexibility with regards to designing the system. For example, confidentiality can be achieved either by using the private data mechanisms, or by creating a separate channel for each pair of communicating organizations.

B. Subsystem Architecture

First, we need to define the high level workflow we expect the officers to perform. A diagram of the workflow can be seen in Fig. 3.

An officer of PIU *A* submits a new PNR data request, officer of PIU *B* responds: either positively (with the requested data) or negatively (with, e.g., a reason for rejection). Both officers need to be able to list all the PNR requests where their organization is either the requester or respondee.

After evaluating various possibilities, we have arrived at the the overall architecture shown in Fig. 4.

The user interacts with a web application programming interface (HTTP API) server, which handles high level user

operations, such as “I want to respond to a request with data”. The HTTP API server translates these operations into more specific operations in a repository layer. The chaincode (smart contract) deployed on a Fabric channel implements the repository interface and translates its operations into reads/writes into public ledger and private data respectively.

Decoupling the actual user operations from chaincode implementation has several benefits. Firstly, the HTTP API becomes the public interface used by RMT to integrate with the blockchain subsystem. This interface is continuously being refined based on user feedback and RMT’s technical requirements; many of these changes can be done without affecting the chaincode implementation. Chaincode changes are disruptive, as even the testing blockchain network needs to be updated or redeployed from scratch. Conversely, work done on the chaincode can often be invisible to RMT and its interaction with the blockchain subsystem. Finally, for development purposes, we can provide technical partners working on RMT with an HTTP API server that uses a repository implementation which does not require a functioning blockchain network, and can be easily (re)deployed locally on developer machines.

Although the HTTP API server validates incoming user requests (e.g., user cannot respond to a request originating in their organization), additional validation is done in the chaincode. While this results in duplication of all the validation rules, we consider this a worthwhile tradeoff to ensure internal consistency of data residing on the blockchain. For the system to truly provide the required guarantees, a malicious actor bypassing the HTTP API and calling the chaincode directly must not be able to change the data in any way that would not also be possible by calling one of the HTTP API endpoints.

C. Data Confidentiality and Access Control

At this stage of the research project, we currently store all the data in the public ledger². As we are only a single part of a larger system, collaborating with several partners, we have prioritized getting a working end-to-end solution that provides them and end users alike with an idea of how the system will work over utilizing the blockchain capabilities to the fullest from the start.

Within a Fabric network, there are two techniques how to provide data confidentiality. One is to create a channel for each pair of participating organizations, as shown in Fig. 5. While this is simple from a development point of view — there is no need to decide what goes into public ledger and what into private data, smaller HTTP API required to call from the chaincode — it is quite complex to manage and operate. Each new organization joining the network results in needing to create $N - 1$ channels with correct configuration. Similarly, it is limiting in the way that giving a third party access to the ledger (i.e., allowing them access to the channel) for auditing

²We do not (and will not) have access to any actual user data, all the data is artificial for development and testing purposes. Likewise, the ledger is public in context of the Fabric channel, but the network itself is permissioned and isolated.

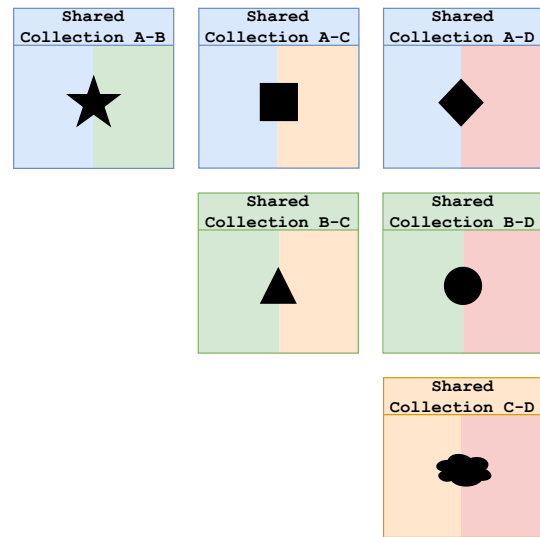


Fig. 5. Mapping of various PNR data requests when using pairwise channels/collections.

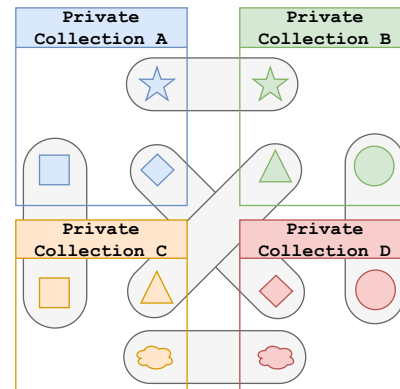


Fig. 6. Mapping of various PNR data requests between organizational private collections.

purposes also necessarily reveals the sensitive personal data. Interaction with the blockchain network from the HTTP API

The second approach is to use *Private Data*. This is a feature of Hyperledger Fabric which allows organizations to selectively share data with a subset of channel participants. The confidential data are exchanged through a gossip side-channel, and the broadcast transaction will contain only hashes of this data. There can be any number of private data collections within a channel, with each collection having its own access control. Each organization, by default, has a corresponding implicit collection readable only by them. Private data collections are a built-in form of off-chain storage, internally implemented using LevelDB or CouchDB.

Noteworthy is the fact that any organization can write to any collection, assuming the endorsement policy for that transaction is satisfied, without being a member of that collection. This enables various approaches of using private data collections. We could create a private collection for each pair of organizations, accessible by both of them. This is analogous

to the multiple channel approach we mentioned previously, and suffers from the same scalability/management complications. A more suitable solution for our use case is to use the implicit per-organization collections. Every change of private data in the chaincode will result in two writes; one into the organization's own collection, and one into the remote peer's (to which the organization has no read access), visualized in Fig. 6. With this approach, a new organization joining the channel results in one new collection being created—created automatically by Fabric itself, with no intervention by system administrators.

An additional benefit of this approach is that it greatly simplifies the work needed to be done on the HTTP API server when interacting with the blockchain network. Consider the scenario of a PIU officer wanting to see all the PNR requests relevant to their unit; that is, those where the unit is either the requester or the requestee. In the separate channel/pairwise collection cases, the server must: a) get a list of all the active *and historical* (i.e., those who have previously participated but left the network) organizations; b) determine the list of channels/collections it needs to access; c) query the channels/collections for data; and d) merge all the data into a final response. The need to be able to access even historical data means some organizations might need to belong to a number of channels/collections even as a single participant. Conversely, with our proposed approach, the requested data are exactly the contents of the organization's private collection, reducing the work required into a single step.

D. Data Protection Safeguards

The PNR Directive states the following with regards to data retention [2]: a) data must be depersonalized after 6 months; b) data may be re-personalized only under strict conditions; and c) data must be deleted after 5 years. It further defines *depersonalize through masking out of data elements* as “means to render those data elements which could serve to identify directly the data subject invisible to a user”.

Depersonalization, therefore, requires understanding the data to be able to selectively remove parts of it. We have intentionally designed the blockchain subsystem with the assumption that both the request contents, as well the PNR data response, are opaque “blobs” of data to us; we merely ensure their secure, traceable exchange.

In other words, there is no reason to keep the data in the system after the exchange has been completed; the system has served its purpose. The HTTP API allows the RMT instance to confirm the receipt of transferred data after each step of the whole process, i.e., after a new request for that instance, or a response for a request originating from that instance, has been received. As part of this confirmation process for a finalized request, we can delete the PNR request/response blobs, leaving behind only metadata about the exchange and hashes of the blobs. We should thus be able to delete the data much sooner than the required 5 years, while anyone possessing the blob can still verify the transfer by comparing the blob's hash against the one in the blockchain.

This handles the “happy path”, where a request receives a response and reaches its finalized state. If a request was created, but never received a response, it will only contain the request blob. It is still an open question that we will need to discuss with our legal partners in the project whether the request blob also falls under these data protection rules, or any other relevant data privacy regulations. Nonetheless, that still leaves the possibility of a response being sent but never confirmed by the requester; a case where the data definitely needs to be deleted.

The fabric has a built-in feature that allows limiting the lifetime of private data. Unfortunately, the lifetime is specified in number of blocks, which is an unreliable unit of time in a network where blocks get created on demand (in contrast with networks such as Bitcoin that aim to have a new block roughly every 10 minutes).

Our proposed solution is to have a periodic job running on the HTTP API server that would delete the blobs from all requests in the organization's private collection which have been created earlier than some time period, e.g., 3 months. This is much more aggressive than the required 5 years, based on the assumption that if an exchange has not been finalized within such a time period, the chances of it happening are slim. We might make the time period configurable, and let the particular organization decide how aggressively they want to delete such lingering data, but the maximum allowed period must not be longer than 5 years.

E. Security Threat Analysis

In designing a blockchain-based system for securing PNR data exchange, we conducted a short security threat analysis, focusing on several categories:

1) Data Confidentiality and Integrity Threats:

- **Eavesdropping and Data Interception:** Unauthorized access to PNR data during transmission between PIUs can lead to breaches of privacy. To mitigate this, we are using Hyperledger Fabric, which utilizes TLS to ensure confidentiality of communication between network participants.
- **Internal Threats:** Insiders with malicious intent can misuse their access privileges to leak sensitive data. This can not be mitigated purely by technological means (e.g., an officer can write the sensitive data down on paper). Within the system, role-based access control limits the access of unauthorized personnel to sensitive data.
- **Data Tampering:** Modification of the PNR data is possible only through chaincode deployed within the network. Deployment of a new chaincode has to be approved by a sufficient number of network participants, depending on the configured policy. Given that each organization is a single participant, no matter the number of peer or orderer nodes they operate, an attacker would have to compromise a number of PIU units to deploy their malicious code. Even if they managed to do this, records of the data being changed (originating organization and the identity within, as well as the calling of a different,

suspicious chaincode) would be present in the immutable ledger for all to audit and verify.

- Impersonation and Forged Data: For example, an illegitimate PIU can inject false PNR data into the system. Public Key Infrastructure (PKI) is employed to authenticate the identities of all nodes in the blockchain network and establish the identity of each PIU, meaning only trusted, authenticated units can participate in the network (after they are admitted with the majority of participants into the network).

2) Availability Threats:

- Denial of Service (DoS) Attack: A DoS attack could target the network to disrupt the availability of the PNR data exchange service. The blockchain network (as a decentralized system) is designed with resilience and redundancy to ensure availability.
- Network Partitioning: In the event of network failures, the network could become partitioned, leading to a risk of inconsistent data states. The consensus mechanism in Hyperledger Fabric is designed to handle such inconsistencies and ensure a consistent agreement on the state of the network.

3) External Threats:

- Malware: External attackers may deploy malware to gain access to or disrupt the PNR data and the system's infrastructure. General endpoint security measures, including firewalls, logs, and IDS should be employed to protect the organization's infrastructure.

The security threat analysis highlights various potential threats to the blockchain-based PNR data exchange system and the countermeasures implemented to address them. By leveraging Hyperledger Fabric's advanced security features and implementing robust PKI for identity verification, the system ensures protection against most threats that could occur.

F. System Performance and Scalability

At this stage of the research project, we have not yet conducted comprehensive performance measurements; such work is planned for a later stage. As part of our future work, we intend to conduct methodical performance evaluations of our solution. These assessments will likely utilize specialized tools designed for blockchain performance analysis, such as Hyperledger Caliper [33], to provide quantitative insights into the system's capabilities and limitations. However, the work described in subsections IV-B, IV-C and IV-D was done with scalability in mind.

There are two main variables how the system can scale: a) the number of participating organizations (PIUs) and the amount of PNR data being exchanged between these participants. It is worth explicitly pointing out that the amount of PNR data being collected — which is by far the largest number and going to keep increasing with time — has no effect on the system and its performance, as this data does not enter the system in the first place.

Given that the system is being developed as part of an EU research project, in the context of EU regulations and legislation, we can expect the number of participating organizations to approximately match the number of EU member states. Although technically nothing prevents a friendly non-EU member state from participating (ignoring any non-technical agreements that are out of the scope of this work), given the permissioned nature of the system and the need for approval from current participants, we can reasonably expect that the growth in this direction will be slow, if any. Therefore, we do not consider scalability in this direction to be of crucial importance. Nonetheless, we have taken this into account and made decisions (such as the design of private collections in subsection IV-C) to be efficient even in this manner.

The other factor—the number of PNR data requests—is more important, as it almost directly translates to possible limitations such as the amount of required storage space. The data structure representing a PNR request (and the associated response during the whole lifecycle of the exchange) in the chaincode can be divided into two parts: the data being exchanged (request and response) and the necessary metadata (status, timestamps, etc.). Of these, the bulk of the total size is the data itself. For bookkeeping purposes, it is not necessary to keep the data around once the exchange has been finalized; in fact, as discussed in subsection IV-D, proactively deleting data as soon as possible is desirable from a privacy protection standpoint. The blockchain subsystem is not intended to be a long-term storage of the PNR data, only a transaction ledger.

Despite the absence of empirical performance data for our specific implementation, it is worth noting that previous studies, such as those by Nasir et al. [34] and Gorenflo et al. [35], suggest that the Hyperledger ecosystem generally exhibits comparable execution times, latency, and scalability to well-established blockchain platforms like Ethereum.

V. CONCLUSION

In this work, we have explored the challenges and solutions related to the secure and tamper-proof exchange of PNR using blockchain technology. The EU data privacy and retention regulations require a robust system to ensure confidentiality, integrity, and traceability of sensitive personal information. Our proposed solution leverages Hyperledger Fabric to create a decentralized, permissioned blockchain network that addresses these requirements. We have detailed the design and implementation of our blockchain subsystem within the TENACITY project, which integrates various crime prevention tools. Using private data collections and off-chain storage mechanisms in Hyperledger Fabric allows for secure data exchange while maintaining compliance with EU directives. Our architecture ensures that only authorized parties can access the data, and all actions are recorded in a non-repudiable manner, thereby enhancing the auditability and accountability of the system. By implementing an HTTP API server that interfaces with the blockchain network, we streamline user operations. This separation of concerns between the HTTP API and the

blockchain backend also simplifies development and maintenance, allowing flexibility in adapting to user and technical requirements. While our current implementation focuses on creating a functional end-to-end solution, future work will address the optimization of data handling and further integration with other components of the TENACITY toolbox. In conclusion, our blockchain-based system for PNR data exchange represents an advancement in the secure handling of sensitive information. By utilizing Hyperledger Fabric, we have demonstrated the feasibility and effectiveness of blockchain technology in enhancing the security and transparency of PNR exchanges between PIUs as a viable alternative to SIENA.

ACKNOWLEDGMENT

The development of the blockchain subsystem, as well as this article, is funded under the HORIZON-CL3-2021-FCT-01-01 project entitled *TENACITY: Travel Intelligence Against Crime and Terrorism*. The authors are grateful for the support and collaboration of the entire TENACITY consortium.

REFERENCES

- [1] "Council Directive 2004/82/EC of 29 April 2004 on the obligation of carriers to communicate passenger data," *Official Journal of the European Union*, vol. L 261, pp. 24–27, Apr. 2004.
- [2] "Directive (EU) 2016/681 of the European Parliament and of the Council of 27 April 2016 on the use of passenger name record (PNR) data for the prevention, detection, investigation and prosecution of terrorist offences and serious crime," *Official Journal of the European Union*, vol. L 119, pp. 132–149, Apr. 2016.
- [3] European Council, *Passenger data*, accessed 20th May 2024, Mar. 2024. [Online]. Available: <https://www.consilium.europa.eu/en/policies/fight-against-terrorism/passenger-name-record/>.
- [4] European Council, *Policies: Law enforcement cooperation, Passenger data*, accessed 20th May 2024, 2023. [Online]. Available: <https://www.consilium.europa.eu/en/policies/fight-against-terrorism/passenger-name-record/>.
- [5] K. Kozak, "Algorithmic governance, code as law, and the blockchain common: Power relations in the blockchain-based society," *Frontiers in Blockchain*, vol. 6, 2023, ISSN: 2624-7852. DOI: 10.3389/fbloc.2023.1109544. [Online]. Available: <https://www.frontiersin.org/journals/blockchain/articles/10.3389/fbloc.2023.1109544>.
- [6] *EBSI*, accessed 20th May 2024, Mar. 2024. [Online]. Available: <https://ec.europa.eu/digital-building-blocks/sites/display/EBSI/Home>.
- [7] The Expert Group on Blockchain Ethics (EGBE), *Ethical guidelines for blockchain systems*, 2024. [Online]. Available: https://ec.europa.eu/digital-building-blocks/sites/download/attachments/674507497/Ethical_Guidelines_for_Blockchain_Systems_compressed.pdf.
- [8] M. Belotti, N. Božić, G. Pujolle, and S. Secci, "A Vademecum on Blockchain Technologies: When, Which, and How," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 4, pp. 3796–3838, 2019. DOI: 10.1109/COMST.2019.2928178.
- [9] M. Staderini, E. Schiavone, and A. Bondavalli, "A Requirements-Driven Methodology for the Proper Selection and Configuration of Blockchains," in *2018 IEEE 37th Symposium on Reliable Distributed Systems (SRDS)*, 2018, pp. 201–206. DOI: 10.1109/SRDS.2018.00031.
- [10] K. Wüst and A. Gervais, "Do you Need a Blockchain?" In *2018 Crypto Valley Conference on Blockchain Technology (CVCBT)*, 2018, pp. 45–54. DOI: 10.1109/CVCBT.2018.00011.
- [11] Hyperledger Fabric Community, *Hyperledger Fabric Documentation*, 2023. [Online]. Available: <https://hyperledger-fabric.readthedocs.io/en/latest/>.
- [12] Hyperledger Besu Team, *Hyperledger Besu*, 2024. [Online]. Available: <https://besu.hyperledger.org>.
- [13] Quorum Development Team, *Quorum Documentation*, 2024. [Online]. Available: <https://docs.goquorum.consensys.io>.
- [14] Corda Development Team, *Corda Documentation*, 2024. [Online]. Available: <https://docs.r3.com/en/platform/corda.html>.
- [15] Hyperledger Fabric Community, *Hyperledger Fabric Private Data Collections*, 2023. [Online]. Available: <https://hyperledger-fabric.readthedocs.io/en/latest/private-data/private-data.html>.
- [16] Tessera Development Team, *Tessera Documentation*, 2024. [Online]. Available: <https://docs.tessera.consensys.io>.
- [17] G. Zyskind, *Secret Network: A Privacy-Preserving Secret Contract & dApp Platform*, en, 2024. [Online]. Available: <https://srt.network/graypaper> (visited on 04/20/2024).
- [18] D. Song, *The Oasis Blockchain Platform*, en, Jun. 2020. [Online]. Available: https://assets.website-files.com/5f59478e350b91447863f593/628ba74a9aee37587419cf65_20200623%20The%20Oasis%20Blockchain%20Platform.pdf.
- [19] H. Yin, S. Zhou, and J. Jiang, *Phala Network: A Secure Decentralized Cloud Computing Network Based on Polkadot*, en, 2022. [Online]. Available: <https://api.semanticscholar.org/CorpusID:233305130>.
- [20] S. Yadav and N. Tiwari, "Privacy preserving data sharing method for social media platforms," *PLOS ONE*, vol. 18, no. 1, pp. 1–20, Jan. 2023. DOI: 10.1371/journal.pone.0280182. [Online]. Available: <https://doi.org/10.1371/journal.pone.0280182>.
- [21] M. Žiška, "Privacy preserving smart-contract platforms and e-voting," Supervisor Ing. Martin Perešini, Master's thesis, Brno University of Technology, Faculty of Information Technology, 2024. [Online]. Available: <https://www.vut.cz/en/students/final-thesis/detail/155970>.
- [22] J. H. Beinke, C. Fite, and F. Teuteberg, "Towards a stakeholder-oriented blockchain-based architecture for electronic health records: Design science research study," *Journal of Medical Internet Research*, vol. 21, no. 10, e13585, Oct. 2019, ISSN: 1438-8871. DOI: 10.2196/13585. [Online]. Available: <http://dx.doi.org/10.2196/13585>.
- [23] M. Antwi, A. Adnane, F. Ahmad, R. Hussain, M. Habibur Rehman, and C. A. Kerrache, "The case of Hyperledger Fabric as a blockchain solution for healthcare applications," *Blockchain: Research and Applications*, vol. 2, no. 1, p. 100012, 2021, ISSN: 2096-7209. DOI: <https://doi.org/10.1016/j.bcr.2021.100012>. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S2096720921000075>.
- [24] L. Van Hoye, T. Wauters, F. De Turck, and B. Volckaert, "Trustful ad hoc cross-organizational data exchanges based on the hyperledger fabric framework," *International Journal of Network Management*, vol. 30, no. 6, e2131, 2020, e2131 nem.2131. DOI: <https://doi.org/10.1002/nem.2131>. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/nem.2131>. [Online]. Available: <https://onlinelibrary.wiley.com/doi/abs/10.1002/nem.2131>.
- [25] C. Schaefer and C. Edman, "Transparent logging with hyperledger fabric," in *2019 IEEE International Conference on Blockchain and Cryptocurrency (ICBC)*, 2019, pp. 65–69. DOI: 10.1109/BLOC.2019.8751339.

- [26] S. Y. Popov, *Iota – the tangle*, 2018. [Online]. Available: https://assets.ctfassets.net/r1dr6vzfxhev/2t4uxvsIqk0EUau6g2sw0g/45eae33637ca92f85dd9f4a3a218e1ec/iota1_4_3.pdf.
- [27] C.-L. Chen, J. Yang, W.-J. Tsaur, W. Weng, C.-M. Wu, and X. Wei, "Enterprise data sharing with privacy-preserved based on hyperledger fabric blockchain in iiot's application," *Sensors*, vol. 22, no. 3, 2022, ISSN: 1424-8220. DOI: 10.3390/s22031146. [Online]. Available: <https://www.mdpi.com/1424-8220/22/3/1146>.
- [28] Passenger Information Unit CZ, Police of the Czech Republic, *PIU CZ – Guide for airlines, PNR data requirements*, accessed 20th May 2024, Aug. 2019. [Online]. Available: <https://www.policie.cz/soubor/piu-cz-guide-for-airlines-august-2019-pdf.aspx>.
- [29] International Air Transport Association, *Passenger and Airport Data Interchange Standards EDIFACT Implementation Guide (version 21.1)*, accessed 20th May 2024, 2021. [Online]. Available: https://www.iata.org/contentassets/18a5fdb2dc144d619a8c10dc1472ae80/pnrgov-edifact-implementation-guide_21_1.pdf.
- [30] International Air Transport Association, *Passenger and Airport Data Interchange Standards XML Implementation Guide (version 16.1)*, accessed 20th May 2024, 2016. [Online]. Available: https://www.iata.org/contentassets/18a5fdb2dc144d619a8c10dc1472ae80/pnrgov20xml20implementation20guide2016_1.pdf.
- [31] Europol, *Secure Information Exchange Network Application (SIENA)*, accessed 20th May 2024, Jun. 2022. [Online]. Available: <https://www.europol.europa.eu/operations-services-and-innovation/services-support/information-exchange/secure-information-exchange-network-application-siena>.
- [32] B. Valentin, L. Gale, H. Boulahya, *et al.*, "BLENDED - Using Blockchain and Deep Learning for Space Data Processing," in *Big Data from Space (BiDS'21)*, May 2021.
- [33] Hyperledger Caliper Team, *Bitcoin network simulator*, online, accessed 25th June 2024, 2024. [Online]. Available: <https://github.com/hyperledger/caliper>.
- [34] Q. Nasir, I. A. Qasse, M. Abu Talib, and A. B. Nassif, "Performance analysis of hyperledger fabric platforms," *Security and Communication Networks*, vol. 2018, no. 1, p. 3976093, 2018. DOI: <https://doi.org/10.1155/2018/3976093>. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1155/2018/3976093>. [Online]. Available: <https://onlinelibrary.wiley.com/doi/abs/10.1155/2018/3976093>.
- [35] C. Gorenflo, S. Lee, L. Golab, and S. Keshav, "Fastfabric: Scaling hyperledger fabric to 20000 transactions per second," *International Journal of Network Management*, vol. 30, no. 5, e2099, 2020, e2099 nem.2099. DOI: <https://doi.org/10.1002/nem.2099>. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/nem.2099>. [Online]. Available: <https://onlinelibrary.wiley.com/doi/abs/10.1002/nem.2099>.