Challenges and Opportunities of Blockchain for Auditable Processes in the Healthcare Sector

Chapter · September 2022

DOI: 10.1007/978-3-031-16168-1_5

3 authors, including:

Walid Fdhila
SBA Research
50 PUBLICATIONS **659** CITATIONS

SEE PROFILE

Challenges and Opportunities of Blockchain for Auditable Processes in the Healthcare Sector*

Walid Fdhila[1,2], Nicholas Stifter[1,2], and Aljosha Judmayer[1,2]

[1] SBA Research
[2] University of Vienna

Abstract. Blockchain technologies (BT) promise to offer exciting research directions for improving various aspects of business processes, in particular in cross-organizational settings where participants do not fully trust each other. However, while blockchain may readily provide transparency and immutability for the processes recorded on a shared ledger, these very characteristics can be problematic in regard to privacy and data protection requirements. In this paper, we address the challenges and opportunities of using BT to secure distributed processes where participants may have an incentive to make false claims or subvert pre-agreed compliance rules in their private processes. Specifically, our analysis is based on a real-world use case, namely how BT can secure (privacy preserving) commitments to processing steps that facilitate federated machine learning (FL) in the healthcare sector. Thereby, an immutable audit trail is created that can be used to detect deviations in retrospect. Hereby, we place a particular focus on the management of patient consent for accessing their data in FL. Our approach draws inspiration from the domain of Self-Sovereign Identity (SSI) where BT is also relied upon to enable the creation and management of decentralized identifiers while focusing on data minimization. The results of our work are not constrained to the particular use case and can be applicable to other emerging research areas of BPM, such as federated process mining.

Keywords: Blockchain, Healthcare, Consent management, Business process

1 Introduction

There is a dichotomy between user privacy/data protection requirements and the construction of a blockchain as a transparent and verifiable immutable ledger of transactions. On the one hand, BT promise to offer compelling characteristics and properties that can be leveraged, e.g., tamper resistance, high reliability, openness, and distributed or even decentralized trust [26,18]. For instance, they can be used to realize global data sharing and data traceability systems, where

* This is a preprint. The final authenticated version is available online at `https://doi.org/10.1007/978-3-031-16168-1_5`

these advantages make it possible to build larger scale, higher quality, and auditable global decentralized data platforms. On the other hand, the aforementioned properties of BT also present fundamental challenges in respect to ensuring user privacy and confidentiality, as well as enabling the removal of undesirable content or otherwise deleting or changing the recorded transaction history [22]. In some application domains, such as the healthcare sector, the ability to both withhold and even delete data due to privacy and regulatory requirements, e.g. the General Data Protection Regulation (GDPR), constitutes a necessity for compliant systems. At first, it would appear that in such cases incorporating BT is not an ideal approach. However, with careful design considerations the advantages of BT can be leveraged while avoiding these privacy issues.

In this paper, we highlight the potential utilization of BT in cross organizational business processes with untrusted parties *where ensuring data privacy and compliance constitutes a necessity* by presenting and analyzing a real-world scenario from the healthcare domain. The use case deals with the management of patient consents in FL. It outlines how access control and audit logs can be implemented through BT in a setting where privacy/confidentiality requirements are high and how commitments in the audit log can be used as a deterrent for misbehavior, even if compliance can not be fully verified automatically. The solution not only showcases how BT may be employed for various aspects of BPM where similar trust-issues could arise, e.g, in federated process mining, it also illustrates how BT can be integrated into legacy systems and processes without full digitization. We hereby bridge an important theory-practice gap in regard to novel proposals leveraging BT and an integration into existing information systems. Further, we show how new paradigms and approaches to identity management in the form of SSI may also be leveraged for the particular use case.

The remainder of this paper is structured as follows: Section 2 contextualizes the research problem and covers related work. Section 3 presents a detailed description of our use case scenario and outlines both, the main requirements, as well as the associated threats, while Section 4 covers system design details. Finally, Section 5 highlights future research challenges and insights gained.

2 Background and Related Work

This section contextualizes the addressed research problem and highlights related work, as well as open challenges, in regard to employing BT in the context of BPM and FL with a focus on privacy and confidentiality.

2.1 Blockchain in Business Process Management

In the field of BPM, the compelling characteristics of BT have garnered interest, in particular in regard to supporting and securing *cross-organizational* processes where involved parties may not fully *trust* each other [26,18,12,7]. Other application areas may be the provision of (immutable) audit trails [25,4,1]. The properties of the underlying blockchain data structure as an immutable totally ordered log of transactions has also been used for *process mining* [15,20,8,14].

It is expected that BT will fundamentally shift how organizations manage their business processes within their network, thereby opening up new challenges and research directions [18]. One such challenge in BPM is the aforementioned dichotomy between privacy/confidentiality and transparency when integrating BT. In cross-organizational settings, where sensitive data is involved, it can be necessary for certain data and processes to remain private, rendering it difficult to verify if they were performed correctly. Distributed compliance checking is able to capture if processes deviate from pre-agreed rules [16], however some misbehavior, such as utilizing healthcare data in private processes for which no consent was given, can evade such detection as the public part of the process may still appear conformant. When relying on BT, the results of private processes and other data that is extraneous to the blockchain must be fed into the system by a so called *oracle* [5]. As the name implies, the oracle[3] may not necessarily be truthful or provide correct results, requiring a certain degree of trust. A tangible example for an oracle is a weather monitoring station that feeds temperature data to a smart contract. [5] investigate how to ensure data confidentiality during business process execution on blockchain even in the presence of an untrusted oracle. However, the solution they propose requires a trusted setup/intermediary and relies on homomorphic encryption schemes for more complex patterns, rendering a practically feasible application limited to specific use cases.

2.2 Applications of Blockchain Technology for FL in Healthcare

A pressing problem in training artificial intelligence (AI) and machine learning (ML) models in healthcare is that the required data, which is usually distributed over various hospitals that may even be located in different jurisdictions, needs to be globally accessible for the training phase, e.g. in a central cloud. This challenge is further compounded by the increasing legal (e.g., GDPR[4],) and technical demands on data providers. In light of enormous data leaks and controversy surrounding how third parties protect data, the public opinion, as well as the patient trust, demand secure ways of storing and processing their data.

Unlike existing approaches, where the ML algorithm runs on data aggregated from different healthcare providers, FL [17] could help address these legal and privacy issues by enabling an ML algorithm to be executed locally at each data provider's site, and the output trained models are subsequently collected and aggregated. This avoids insecure client-cloud and inter-cloud communication, and can ensure that all data remains within (legally and technically) the data provider infrastructures. However, because data is kept and processed locally, external parties cannot readily ascertain from the output results that the process was carried out correctly. In particular, this approach relies on the assumption that all actors involved in the FL are trusted and behave honestly according to the established protocol, meaning that the parties are required to provide correct results and rely on eligible and untampered training data.

[3] Referred to by Weber et al. in [26] as a *trigger*.
[4] Cf. Art. 17 GDPR Right to erasure `https://gdpr-info.eu/art-17-gdpr/`

Experience has shown that such strong trust assumptions often do not hold in practice [24], and thus, a clear threat model that accounts for both privacy and security threats needs to be elaborated. For example, in the context of a study on COVID-19 vaccine efficiency, a healthcare data provider could attempt to bias the output results in order to render the results of a specific vaccine more favourable, or illicitly include data for which they do not have patient consent. Hence, some form of manual or automatic auditing process is required to verify the integrity of the results and that no data manipulation was carried out during the learning process. To address this problem, prior art has both, considered Byzantine resistant *aggregation* of ML outputs that rely on advanced cryptographic schemes such as MPC, e.g. [13], and proposed approaches that seek to render the process more accountable, e.g. through use of blockchain [27,21,19]. In the context of improving clinical research quality, in particular securing the auditing process, BT could offer desirable characteristics [3] while retaining the decentralization afforded from FL as well as remaining largely compatible with legacy system designs. However, as the application of BT in electronic health is still a young concept careful consideration and sufficient risk analysis must be conducted [22] to avoid security threats and flawed system architectures.

3 Use Case: Auditable Consent Management for Federated Machine Learning in Healthcare

This section describes a use case scenario from the healthcare domain, captures the main requirements, and outlines the associated threats. The use case is part of the European H2020 Featurecloud project[5], which provides a privacy preserving solution for FL of healthcare data.

3.1 Use Case Description, Roles and Workflow

In the healthcare sector, FL is a distributed and privacy-preserving learning process that involves several medical data providers who collaboratively train a model. Data of different providers are locally trained and the output models are aggregated either i) centrally by a coordinator, or ii) in a distributed manner, e.g., using cryptographic techniques such as secure multi-party computation (SMPC). Fig. 1 depicts a simplified BPMN collaboration diagram that captures the main steps and necessary interaction flows for a FL study. In this example, a research entity (i.e., *the project coordinator*) wants to conduct a study on breast cancer survival. Therefore, gene expression data from patients are required to learn a model that predicts tumor recurrence. The project coordinator would select and invite participating hospitals (i.e., *a participant*) where each has a local project manager (LPM) assigned to the project. The invitation includes meta-data about the scope of the study and participation requirements (e.g., conditions on data). Upon acceptance to join the study, a participant should perform a data discovery check to compute the amount of consented data relevant to the study, upon which its participation in the study is either confirmed or

[5] https://featurecloud.eu/

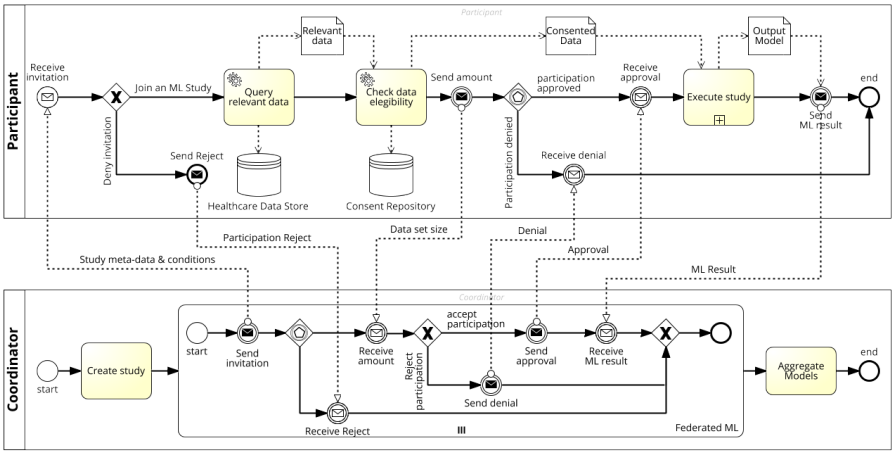

Fig. 1. Collaboration diagram: Federated Machine Learning in Healthcare

denied. If the conditions are met, the project coordinator sends a participation approval that includes a data analysis (ML) application that can be executed locally by the participants (i.e., data providers). The latter are also responsible for ensuring the pre-processing pipeline (e.g., format conversion, data standardisation). After collecting all ML output models, the project coordinator aggregates and eventually publishes the results.

3.2 Requirement Analysis

Although, in principle, federating the ML process across participants solves particular privacy and legal issues encountered in centralized learning (as the data remains within its jurisdiction), in practice it raises additional fundamental challenges that need to be addressed. Particularly, in adversarial or cooperative settings, ensuring the integrity and compliance of the local training data, or the reliability of the participating actors can be challenging. Next, we will analyse some of the technical, data protection and privacy requirements for an auditable and compliant federated ML process in the healthcare sector (cf. Table 1).
For example, the legal requirement $R1$ reflects the text of Article 7 (3) of the GDPR, in which patients shall have the right to give, update or revoke consents for future use of their data: *"The data subject shall have the right to withdraw his or her consent at any time"*. Furthermore, during an audit, participants must be able to prove that they were entitled to use the data when the study was performed. This proof of consent ($R2$) is inline with the GDPR (see Article 7 (1)): *"Where processing is based on consent, the controller shall be able to demonstrate that the data subject has consented to processing of his or her personal data"*. Note that consent revocation means that the information that the patient has given consent should be deleted ($R3$). This deletion is not retroactive, i.e., any study which took place before the revocation will not become illegal on withdrawal. Consent can be either digital or a digitized signed paper form. Besides,

Table 1. Requirements

ID	Type	Description
R1	Legal	Revoking, modifying of consent must be possible
R2	Legal	Proof of consent must be possible
R3	Legal, privacy	Consent must be deleted if revoked
R4	Technical, Procedural	Commitment to used data items (and the related consents)
R5	Technical, Procedural	The coordinator must commit results collection & aggregation
R6	Technical, Procedural	Random selection of sites/hospitals for audit
R7	Technical, Procedural	Studies & commitments must be audited
R8	Technical, Procedural	Auditors must commit the results of the audit
R9	Privacy, Technical	The used commitments must not reveal the original input data (hiding)
R10	Technical	Commitments must be cryptographically signed by respective parties

they can be i) static, i.e., are given once for all future studies, limited with an expiry date, or (ii) dynamic, i.e., are given on the fly for each study. Note that consents are often locally managed by participants, but in an ideal scenario, may be managed by patients (cf. Section 4).

To ensure reliable audits, the tasks executed by the coordinator and each participant should be committed to the blockchain ($R4 - 10$). Commitments can be realized through cryptographic primitives, which allow one to commit to a chosen input value that is not revealed. In $R4$, the participant must commit to the inputs and outputs of the locally executed ML study, as well as the respective consents. These commitments might be checked by a competent authority, i.e., the *auditor*, against the actual data. In combination with the appropriate penalties in case of a wrong doing, this provides a credible threat that deviations are detected and thus probabilistically guarantees the integrity of the FL results and the process compliance with the aforementioned rules (e.g., GDPR).

3.3 Threat Model

In the following, a threat model is elaborated with a focus on threats that influence basic design decisions for securing FL processes and managing consents. The threat modelling follows a mixture of the threat modelling defined in the Microsoft SDL[6] and attack trees[7], where threats are modeled as attack trees with attack goals as roots and alternative ways to achieve that goal as tree branches. The full list of identified threats as well as mitigations can be found in [11]

As participants through their local project managers (LPMs) obviously have access to all patient data for technical reasons, there is no straightforward way to prevent them using this data in an unlawful or unauthorized manner for which the patient has not given its consent to. Therefore, it is out-of-scope of this paper to devise technical means that generally prevent the leakage of patient data in all possible ways, as this would require hospitals to give up at least some control over their IT infrastructure and thus might have other error-prone side effects. Instead we assume, that the participants (i.e., hospitals) do not want to intentionally leak patient data, but they might be willing to influence FL studies

[6] https://www.microsoft.com/en-us/securityengineering/sdl/threatmodeling

[7] https://www.schneier.com/academic/archives/1999/12/attack$_t$rees.html

by including non consented data, fake data, or omit consented data. Moreover, they might want to learn about patient data from other hospitals. In relation to the IT security literature, the participants can thus be considered *covert adversaries* [2], that only act maliciously when the chance of not being detected is high. Therefore, it is important to identify wrongdoing or misbehaviour of the main actors during such a FL process. Based on the threat model, a secure design for trusted FL processes will ensure that only eligible and consented data is used. Table 2 provides a classification of possible threats grouped by actors.
In FL, only the output models are aggregated, which means that it is difficult for an external entity to tell whether the model was in fact trained based on consented, poisoned ($T2$) or non consented ($T1$) data. Because the internal processes (e.g., fetching and preparing data) are not visible to the outside, it is possible for the LPM to introduce fake data, thereby biasing the outcome of the study, e.g., to manipulate the efficiency of a specific COVID-19 vaccine. Inversely, an adversary may also exclude eligible input data ($T3$), e.g., data that negatively influences a recommendation for a specific drug. Similarly, a participant can issue fake consents or manipulate the scope of existing ones ($T4$).
In order to render the audit impractical, a participant may intentionally delete patient data ($T5$) or omit the creation of an audit entry ($T6$), both required for integrity and compliance checks on the conducted studies. Alternatively, due to the possibility of technical failures in the processes assigned to the LPM, it is possible that failures exhibit the same outcome as a malicious project manager would ($T6$). This is in particular the case for omission failures, e.g. data loss of digitized consent forms, but can also occur during the digitization process, e.g. flipping of digits in patient data, or erroneous OCR. Hence, even if we assume a strong system model where the LPM is completely trustworthy, it is necessary to contemplate the possibility of (random) technical failures and their potential impact to the correctness and security of the conducted ML studies.
With respect to the project coordinator, it is possible that the provided ML algorithm embodies malicious code that exfiltrates sensitive patient data ($T8$), e.g., leaking information about the input data in the result. Furthermore, the ML algorithm may be designed in a way that it will always yield the outcome wanted by the coordinator, regardless of the input ($T9$). Alternatively, the coordinator can manipulate the aggregation of the collected models ($T10$).
Finally, one of the legal rights of patients with respect to GDPR is to have their data deleted if requested. This, in general, does not represent a threat, but may influence audit results where the actual data is compared to the commitments (audit records). Therefore it becomes impossible to check whether the corresponding commitment stored for auditing purposes corresponds to the data used for a study. This may influence a possible replay of a study on the input data as some of the data were deleted. The deletion may be requested by the subjects for privacy reasons, or maliciously executed by LPMs.

4 System Design

Based on the requirements gathered in Section 3.2, and the threat model elaborated in Section 3.3, we propose a system design that uses BT to secure FL

Table 2. Classification of Threats

ID	Actor	Class	Threat Description
T1	Participant	Data poisoning	Selects non consented data or data with expired/revoked consent
T2	Participant	Data poisoning	Uses fake data as input to the ML algorithm
T3	Participant	Data poisoning	Excludes eligible data
T4	Participant	Tampering	Issues fake consents or manipulates their scope
T5	Participant	Integrity	Deletes data necessary for audit
T6	Participant	Integrity	Data loss due to hardware issues on-site
T7	Participant	Integrity	Does not create an audit log entry
T8	Coordinator	Privacy	Provides a malicious ML algorithm to Leak data
T9	Coordinator	Maniplulation	Provides a malicious ML algorithm to manipulate the outcome
T10	Coordinator	Tampering	Manipulates the model aggregation
T11	Patient	Audit Integrity	Requests data deletion after a study

processes and improve their auditability. A key challenge for the design is to ensure accountability of the involved actors and prove that only legitimate data was used in the FL process, thereby enabling detect wrong doings such as the use of non consented data, or fake patient data, identities and consents. This will help improve the integrity of FL studies and tackle the data poisoning problem in the presence of Byzantine actors where data protection is an issue.
Another important challenge when dealing with consents and data lies in the difficulty for verifiers to prove that the latter were issued or belong to real patients, and not created by participants to bias a study. Therefore, the solution design should carefully consider how identities are linked to consents and data. So far, most healthcare systems relied on identities that are issued by central authorities (e.g, social security number), which if not managed correctly, may in retrospect become a correlation point across FL executions.
We opted for a design that enables a reliable *post auditing* process, and which supports different identity schemes. The design uses a consensus algorithm and a blockchain as an audit trail (for hashes with high min-entropy input, and signatures of involved parties). It also supports both digital and paper-based processes (e.g., handwritten or digital signatures on consent forms). This allows hospitals with different internal processes to participate in the FL.

Next, we first briefly outline the importance of a framework that deals with governance policies for establishing and maintaining a trusted blockchain network, and managing identities. Then, we discuss different identities schemes that may be used within the framework, with a particular focus on self-sovereign identity, i.e., an emerging approach that can give patients more control over their data and consents. Afterwards, we provide a verifiable approach for managing consents, which will enable auditors verify that i) consents are linked to real patient identities (cf. Section 4.1), and that ii) only consented data and most up-to-date consents are used for an FL study (cf. Section 4.2). Due to space restrictions, Section 4.3 only briefly discusses how to extend some activities in the FL process of Figure 1 with constructs that render their auditability more trustworthy (e.g., using commitments on data inputs). Finally, we provide a short overview on the prototype implementation.

4.1 Governance Framework, DPKI and Identities

Healthcare Trust Framework The trust framework (e.g. European actors from the healthcare sector, i.e. Health ministries) defines what issuers will issue what identity or credentials under which policies. This network of trusted authorities (TA) defines the governance rules for all stakeholders and enables legally binding relationships (e.g., onboarding policies, applicable regulations, governance rules, protocols). Members of the trust framework are also responsible for onboarding new members to FeatureCloud (e.g., hospitals, pharmaceutical companies, test labs, insurances) with specific access rights (e.g., writing to the ledger), and monitoring their compliance with the rules. All accredited members in this trust framework may act as both validator nodes and identity providers. Table 3 gives a summary of the stakeholders' roles in the system.

DPKI and Identities Most approaches on healthcare processes (cf. Section 2) rely on the assumption that identities of all actors involved in such collaborative settings are well defined, and that there is an established public key infrastructure (PKI) that governs the issuance of digital certificates, and enables trustworthy communication and authentication. However, dealing with sensitive data such as in healthcare requires a particular level of protection against data misuse by providing mechanisms and means that improve their privacy and security. Thus, an elaborate design of an identity layer is crucial for trusted infrastructure that secures and links public keys to real identities. Recent developments in this area such as certificate transparency already embrace authenticated data structures as a means of identifying manipulation attempts. For example, electronic health certificates have increasingly been relied upon in an effort to contain the spread of COVID-19 infection. Hereby, the difficulty lies in striking a balance between ensuring an individual's privacy and the ability to verify that the claims made in the certificate are authentic and tied to a particular identity.

Unlike previous identities systems, which used to rely on centralized or federated architectures that have already proven their inefficiency and lack of security and privacy, a new identity approach, i.e. SSI [23,10], has emerged, which promises users control over their data, and ensures individuals are at the center of interactions. SSI often relies on blockchain technologies to record identity information or serve as the basis for a Decentralized public key infrastructures (DPKI) [6]. Such Blockchain-based DPKI infrastructure could, for instance, help provide more robust mechanisms for establishing (and revoking) digital identities that are used for aspects such as access control or rights management. SSI initiatives resulted in two key concepts i) verifiable credentials (VC), and ii) decentralized

Table 3. Roles and responsibilities

Coordinator • Initiates a FeatureCloud study • Selects study participants • Provides ML algorithms	**Participant** (e.g., hospital) • Ensures local data security & privacy • Performs data querying & preparation • Executes & manages federated ML studies • Manages patient identities & consent
Trusted Authorities (e.g., Health ministry) • Define policies & onboarding rules • Establish & governs the trust framework • Register & manage stakeholders Identities	**Patient** • Gives, updates & revokes consents **Auditor** • Audits studies

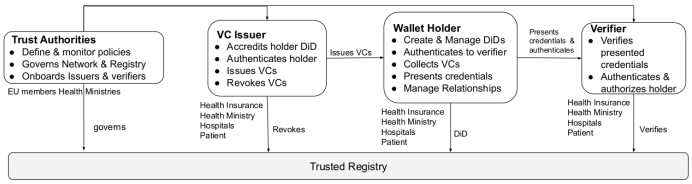

Fig. 2. SSI Actors and Responsibilities

identifiers (DiDs) by the World Wide Web Consortium (W3C):

- *Decentralized Identifiers (DiDs)*: A DID is a globally unique identifier, cryptographically generated (not necessarily registered on a distributed ledger), which points to a DID document (e.g. a JSON-LD object) that specifies cryptographic keying material (e.g., public keys for authentication), verification methods essential for proving ownership of the DID and eventually service endpoints for trustworthy and persistent communication channels.
- *Verifiable Credentials (VCs)*: VCs are tamper-evident identity attributes and assertions about a specific subject issued by an identity provider. In contrast to other types of digital credentials, a relying party (third party service) can check the validity of a VC without having to interact with the issuer (i.e. preventing correlation) (cf. Figure 2). Note that the same stakeholder may play different roles, i.e., issuer, holder or verifier (e.g., a patient issues consents to hospital (holder), but also authenticates within a hospital (verifier)).

Unlike other alternatives such as x.509 certificates or qualified digital signatures, self-sovereign identities can be constructed in a way that is i) decentralized, ii) privacy-preserving (e.g., prevents linkability and supports zero-knowledge proofs), and iii) more secure (e.g., no single point of failure). In a report by the European Union Agency for Cybersecurity (ENISA) [9], an assessment of the potential of SSI technologies and other eID solutions for ensuing secure electronic identification and authentication was provided. It was acknowledged that SSI provides an effective basis for digital identities, which protects the privacy of personal data, but also needs to co-exist/operate with established technologies such as X.509 PKI, OpenID Connect and other identification schemes.
Despite all the capabilities offered by SSI technologies, in practice, it is optimistic to assume that all stakeholders (e.g., hospitals, health insurances) will employ the same identity scheme. Therefore, in the context of FeatureCloud, we recommend using SSI solutions (e.g., DiDs, VCs), but we do not restrict stakeholders from adopting different methods as long as they securely identify and authenticate subjects and protect their privacy. Additionally, it is noteworthy that a binding identity is necessary to prevent malicious participants from biasing output ML models through poison attacks, i.e. using fake identities, consents or data. This means that identities need to be accredited by members of the trust framework or stewards. Furthermore, the identity model must avoid cross-correlating patient data across multiple participants.

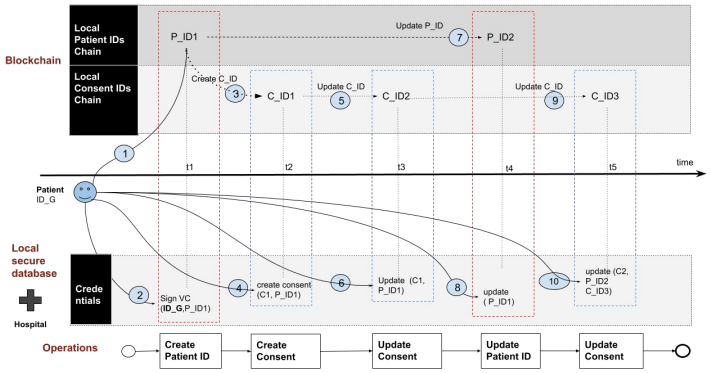

Fig. 3. Process example of consent Commitments

4.2 Verifiable Consent Management Process

Setup and Assumptions It is assumed that the software (e.g., a docker container) provided by the coordinator is trusted. This means that the software does not leak data and is designed to execute as reproducible as possible, i.e., ideally the execution is fully deterministic such that running it with the same inputs results in the exact same output. In the FeatureCloud AI store[8], the docker image of a federated algorithm (FeatureCloud app) is published by a developer, and then certified by FeatureCloud after checking its privacy and performance. In FeatureCloud, both paper and digital consents as well as digital and non-digital identities are supported to prevent discrimination. In the following, we focus solely on the use case where consents are digital or digitized by participants, and patients can use applications to cryptographically sign and manage credentials. We also assume that public representations of cryptographic keying material generated by the patients was signed by a competent authority (e.g., ministry of health), and that there exist one or multiple trusted registries (not necessarily a blockchain) for revoking identities and credentials (cf. Figure 2).
Based on the legal and technical requirements identified in Section 3, a permissioned blockchain can serve as a distributed and immutable audit trail. Access controls (read and write to the ledger) are also defined to restrict access to non-authorized subjects. A smart contract that is deployed and governed by the Trust framework, will be used by authorized entities for registering/updating and revoking local identifiers (e.g., patient identifier within a hospital). Separate smart contracts are also deployed for managing ML studies and consents respectively (i.e., define the rules/logic for commitments).

Patient Identity registration In order to avoid linkability of patient data across participants, each participant must accredit a separate local identifier (PID_l) for each patient. First, the patient has to authenticate himself to the participant using a supported authentication mechanism. Then, for example if SSI is supported, the patient may create a new DiD and registers it within the participant, e.g., by issuing a Verifiable credential of the DiD signed with PID_g.

[8] https://featurecloud.ai/

The local identifier (DiD) is then added by the participant to a trusted registry specific to identified patients. Note that using SSI, only the patient controls the DiD, and can easily rotate the keys (using CRUD operations), while the authority that issued/signed cannot link him to the new PID_l. This is also important because during an audit it proves for an auditor that this PID_l corresponds to an actual patient and prevents generating fake patient/consent identifiers. A patient will use a PID_l to issue or revoke consents to a participant.

Consent Collection As described in Section 3.2, we distinguish two cases with regard to the time at which a consent was given: *i)* upfront, where consents from patients are collected by hospitals upfront in a generic way that is not tailored towards a specific study, i.e., they are usually broader and might be eligible to multiple studies. Here, legal criteria regarding consents in the healthcare system have to be taken into account that might be different for each EU member state. *ii)* on demand, where the hospital actively asks matching patients for consent to use their data in a specific study. In both cases, the consent is never stored in the blockchain or transmitted to another party despite the hospital.

Give consent: using the local identifier, the patient digitally signs and issues the consent to the participant. Ideally, with the right permissions defined in the consent management smart contract (given by the participant), the patient may use his application to commit the consent (CID) to the blockchain (cf. Figure 3). Alternatively, the participant commits the consent on his behalf. The consent itself is stored locally at the hospital site.

Update/revoke consent: the patient informs the hospital that he wants to update or revoke his consent. The updated/revoked consent is signed using PID_l, then issued to the participant. Ideally, the patient uses his wallet to update/revoke the previous commitment after permission is granted by the participant (using the CRUD operations in the smart contract). Operations on the same consent are linked, that an auditor can later verify the consent history to check if it was valid (e.g., expiration date) when the study was executed.

4.3 Federated ML execution

Execute the federated study and commit to input, output and used consents (participants) Each participant executes the study software (e.g., docker container) with the required and consented input data. To facilitate auditability, the participants commit to the patient data used as input to the study, the eventual data transformation operations, the associated consents, and the output results. The commitments must be cryptographically hiding, signed and published on a secure bulletin board, e.g. BT, accessible to all involved parties.

Aggregate all submitted results and signal successful aggregation (coordinator) When all participants have submitted their trained models, i.e., partial results, the coordinator performs the aggregation of the submitted results. After all results have successfully been aggregated and the final output has been computed, the coordinator has to commit to the collected results, as well as to the outcome of the final aggregation. Again, ideally the study is deterministic, and thus the aggregation of the same partial inputs yields exactly the same overall result. In this case, the commitment can be a cryptographic hash

of used inputs and outputs. Otherwise, other techniques such as fuzzy hashing, which enable checking hashes similarity may be employed.

The audit is performed (auditors & participants) In this step, the on-site audits are performed by the auditors. To avoid bias, auditors are randomly selected and assigned to participants. The federated study is recalculated locally on hardware of the auditor, using the patient data provided by the selected participant. In case of a fully deterministic ML setup, the resulting output of the audit's execution should match the study result which was previously submitted and committed by the respective participant. The scope, state and associated signatures of the consents are checked. If the federated study was not designed to be completely deterministic, the entire committed-to input and output, has to be retained to allow for reproducibility during the audit. The auditor re-runs the federated study on his hardware and then determines if the previous output model is sufficiently close to his output (using the appropriate similarity techniques), provided the same input data and training parameters are used.

Although not yet considered in the design, by having the patients additionally committing to the data collected about them e.g., during hospital visits, it becomes possible for the auditor to detect if the participant has manipulated or excluded eligible data. Furthermore, by signing the underlying medical data (hash of the data) in their consents, patients ensure that they really had the described medical treatment / disease. This double checks that the data the participant can use in the studies is correct.

4.4 Implementation

This work has been implemented as part of the European H2020 Featurecloud project, which provides a privacy preserving platform for FL of healthcare data. Part of the framework is an AI Store[9] for FL as an all-in-one platform for biomedical research and other applications. The platform allows to run, develop and publish federated and privacy-preserving machine learning algorithms. All apps are stored as Docker images in a registry where images can be pushed or pulled in accordance with the access rights controlled through an authentication server. On the one hand, each participant needs to run a FeatureCloud controller, which manages the local execution of the ML application. On the other hand, a coordinator controller will be responsible of orchestrating the execution and instructing the participants' controllers to ensure a globally synchronous execution. The PoC is still being tested and improved and has not yet been integrated within the FeatureCloud platform[10]. A Hyperledger Fabric test net has been deployed with few nodes acting as the root of trust (governance framework). In the current version, only x.509 certificates and Idemix identities are supported for authentication as they are inherently supported by Fabric. We are still experimenting with both Hyperledger Indy and Aries to integrate issuance and verification of DIDs and verifiable credentials within Fabric. The implemen-

[9] https://featurecloud.ai/

[10] The source code will be released on gitlab. The code is still in review as it may fall under a temporary NDA agreement

tation also relies on two smart contracts (chaincode) for managing consents and study related data commitments.

5 Insights, Applications, and Future Research Challenges

In this paper we have demonstrated how blockchain can be leveraged in cross-organizational business processes where there exist additional privacy requirements and constraints through a tangible use case. A key problem that we address is how trust in the correctness and adherence to compliance rules within private processes of participants can be increased, as the necessary data for complete verification is not publicly available. We rely on the compelling properties of blockchain to secure commitments to the execution of these private processes that can later be audited. In this regard, the utilization of BT does not *prevent* misbehavior, however it offers *detectability* and non-repudiation as a strong deterrent. Hereby, our solution bridges the current gap between research proposals where verification of private processes is made possible in a privacy preserving manner through MPC or by relying on trusted execution environments (TEEs), and real-world information systems where such designs are currently infeasible or even impossible to deploy. We believe that the herein presented approach presents a practical design pattern for a variety of real-world cross-organizational business processes. Further, our approach can also apply to other related research areas of BPM such as federated process mining. Future research challenges include how the aforementioned advanced cryptographic proof techniques can be effectively integrated into legacy information systems, as well as exploring the potential SSI has to offer in the context of BPM.

Acknowledgments. This research is based upon work partially supported by (1) SBA Research (SBA-K1); SBA Research is a COMET Center within the COMET – Competence Centers for Excellent Technologies Programme and funded by BMK, BMDW, and the federal state of Vienna. The COMET Programme is managed by FFG. (2) the European Union's Horizon 2020 research and innovation programme under grant agreement No 826078 (FeatureCloud) (3) the FFG ICT of the Future project 874019 dIdentity & dApps. (4) the FFG Industrial PhD project 878835 SmartDLP. (5) the Christian-Doppler-Laboratory for Security and Quality Improvement in the Production System Lifecycle;

References

1. Akhtar, A., Shafiq, B., Vaidya, J., Afzal, A., Shamail, S., Rana, O.: Blockchain based auditable access control for distributed business processes. In: International Conference on Distributed Computing Systems (ICDCS). pp. 12–22 (2020)
2. Aumann, Y., Lindell, Y.: Security against covert adversaries: Efficient protocols for realistic adversaries. Journal of Cryptology **23**(2), 281–343 (2010)
3. Benchoufi, M., Ravaud, P.: Blockchain technology for improving clinical research quality. Trials **18**(1), 1–5 (2017)
4. Bonyuet, D.: Overview and impact of blockchain on auditing. International Journal of Digital Accounting Research **20**, 31–43 (2020)

5. Carminati, B., Rondanini, C., Ferrari, E.: Confidential business process execution on blockchain. In: international conference on web services (icws). pp. 58–65 (2018)
6. Christopher, A., et al.: Decentralized public key infrastructure, `https://danubetech.com/download/dpki.pdf`
7. Corradini, F., Marcelletti, A., Morichetta, A., Polini, A., Re, B., Tiezzi, F.: Engineering trustable and auditable choreography-based systems using blockchain. Transactions on Management Information Systems (TMIS) **13**(3), 1–53 (2022)
8. Duchmann, F., Koschmider, A.: Validation of smart contracts using process mining. In: ZEUS. CEUR workshop proceedings. vol. 2339, pp. 13–16 (2019)
9. Evgenia, N., Viktor, P., Marnix, D.: Leveraging the self-sovereign identity (ssi) concept to build trust. Tech. Rep. 10.2824/8646, The European Union Agency for Cybersecurity (ENISA) (2022)
10. Fdhila, W., Stifter, N., Kostal, K., Saglam, C., Sabadello, M.: Methods for decentralized identities: Evaluation and insights. In: BPM: Blockchain and Robotic Process Automation Forum, Rome, Italy. vol. 428, pp. 119–135 (2021)
11. Fenghong, Z., Aljosha, J., Walid, F., Nicholas, S.: D6.2: "model for defining user rights in federated machine learning". Tech. rep., EU H2020 FeatureCloud (2021)
12. Garcia-Garcia, J.A., Sánchez-Gómez, N., Lizcano, D., Escalona, M.J., Wojdyński, T.: Using blockchain to improve collaborative business process management: Systematic literature review. IEEE Access **8**, 142312–142336 (2020)
13. He, L., Karimireddy, S.P., Jaggi, M.: Secure byzantine-robust machine learning. arXiv preprint arXiv:2006.04747 (2020)
14. Hobeck, R., Klinkmüller, C., Bandara, H., Weber, I., van der Aalst, W.M.: Process mining on blockchain data: a case study of augur. In: International conference on business process management. pp. 306–323. Springer (2021)
15. Klinkmüller, C., Ponomarev, A., Tran, A.B., Weber, I., Aalst, W.v.d.: Mining blockchain processes: extracting process mining data from blockchain applications. In: International Conference on Business Process Management. pp. 71–86 (2019)
16. Knuplesch, D., Reichert, M., Pryss, R., Fdhila, W., Rinderle-Ma, S.: Ensuring compliance of distributed and collaborative workflows. In: 9th IEEE International Conference on Collaborative Computing: Networking, Applications and Worksharing. pp. 133–142. IEEE (2013)
17. McMahan, B., Moore, E., Ramage, D., Hampson, S., y Arcas, B.A.: Communication-efficient learning of deep networks from decentralized data. In: Artificial intelligence and statistics. pp. 1273–1282. PMLR (2017)
18. Mendling, J., Weber, I., Aalst, W.V.D., Brocke, J.V., Cabanillas, C., Daniel, F., Debois, S., Ciccio, C.D., Dumas, M., Dustdar, S., et al.: Blockchains for business process management-challenges and opportunities. ACM Transactions on Management Information Systems (TMIS) **9**(1), 1–16 (2018)
19. Mugunthan, V., Rahman, R., Kagal, L.: Blockflow: An accountable and privacy-preserving solution for federated learning. arXiv preprint arXiv:2007.03856 (2020)
20. Mühlberger, R., Bachhofner, S., Ciccio, C.D., García-Bañuelos, L., López-Pintado, O.: Extracting event logs for process mining from data stored on the blockchain. In: International Conference on Business Process Management. pp. 690–703. Springer (2019)
21. Passerat-Palmbach, J., Farnan, T., Miller, R., Gross, M.S., Flannery, H.L., Gleim, B.: A blockchain-orchestrated federated learning architecture for healthcare consortia. arXiv preprint arXiv:1910.12603 (2019)
22. Ploder, C., Spiess, T., Bernsteiner, R., Dilger, T., Weichelt, R.: A risk analysis on blockchain technology usage for electronic health records. Cloud Computing and Data Science pp. 20–35 (2021)

23. Preukschat, A., Reed, D.: Self-sovereign identity. Manning Publications (2021)
24. Schüler, P., Buckley, B.: Re-Engineering clinical trials: Best practices for streamlining the development process. Academic Press (2014)
25. Snow, P., Deery, B., Lu, J., Johnston, D., Kirby, P., Sprague, A.Y., Byington, D.: Business processes secured by immutable audit trails on the blockchain. Brave New Coin (2014)
26. Weber, I., Xu, X., Riveret, R., Governatori, G., Ponomarev, A., Mendling, J.: Untrusted business process monitoring and execution using blockchain. In: International conference on business process management. pp. 329–347. Springer (2016)
27. Weng, J., Weng, J., Zhang, J., Li, M., Zhang, Y., Luo, W.: Deepchain: Auditable and privacy-preserving deep learning with blockchain-based incentive. IEEE Transactions on Dependable and Secure Computing **18**(5), 2438–2455 (2021)