A Case Study for Blockchain in Healthcare:

"MedRec" prototype for electronic health records and medical research data

White Paper

Ariel Ekblaw[*], Asaph Azaria[*], John D. Halamka, MD[†], Andrew Lippman[*]

[*]MIT Media Lab, [†]Beth Israel Deaconess Medical Center

August 2016

Note: The abstract and first three sections of this white paper are drawn from a peer-reviewed, formally accepted paper, published by IEEE.

MedRec: Using Blockchain for Medical Data Access and Permission Management

IEEE Original Authors: Asaph Azaria, Ariel Ekblaw, Thiago Vieira, Andrew Lippman

Abstract

A long-standing focus on compliance has traditionally constrained development of fundamental design changes for Electronic Health Records (EHRs). We now face a critical need for such innovation, as personalization and data science prompt patients to engage in the details of their healthcare and restore agency over their medical data. In this paper, we propose MedRec: a novel, decentralized record management system to handle EHRs, using blockchain technology. Our system gives patients a comprehensive, immutable log and easy access to their medical information across providers and treatment sites. Leveraging unique blockchain properties, MedRec manages authentication, confidentiality, accountability and data sharing—crucial considerations when handling sensitive information. A modular design integrates with providers' existing, local data storage solutions, facilitating interoperability and making our system convenient and adaptable. We incentivize medical stakeholders (researchers, public health authorities, etc.) to participate in the network as blockchain "miners". This provides them with access to aggregate, anonymized data as mining rewards, in return for sustaining and securing the network via Proof of Work. MedRec thus enables the emergence of data economics, supplying big data to empower researchers while engaging patients and providers in the choice to release metadata. The purpose of this paper is to expose, in preparation for field tests, a working prototype through which we analyze and discuss our approach and the potential for blockchain in health IT and research.

1. Introduction

EHRs were never designed to manage multi-institutional, life time medical records. Patients leave data scattered across various organizations as life events take them away from one provider's data silo and into another. In doing so they lose easy access to past data, as the provider, not the patient, generally retains primary stewardship (either through explicit legal means in over 21 states, or through default arrangements in the process of providing care) [1]. Through the HIPAA Privacy Rule, providers can take up to 60 days to respond (not necessarily to comply) to a request for updating or removing a record that was erroneously added [2]. Beyond the time delay, record maintenance can prove quite challenging to initiate as patients are rarely encouraged and seldom enabled to review their full record [1], [2]. Patients thus interact with records in a fractured manner that reflects the nature of how these records are managed.

Interoperability challenges between different provider and hospital systems pose additional barriers to effective data sharing. This lack of coordinated data management and exchange means health records are fragmented, rather than cohesive [3]. Patients and providers may face significant hurdles in initiating data retrieval and sharing due to economic incentives that encourage "health information blocking." A recent ONC report details several examples on this topic, namely health IT developers interfering with the flow of data by charging exorbitant prices for data exchange interfaces [4].

When designing new systems to overcome these barriers, we must prioritize patient agency. Patients benefit from a holistic, transparent picture of their medical history [3]. This proves crucial in establishing trust and continued participation in the medical system, as patients that doubt the confidentiality of their records may abstain from full, honest disclosures or even avoid treatment. In the age of online banking and social media, patients are increasingly willing, able and desirous of managing their data on the web and on the go [3]. However, proposed systems must also recognize that not all provider records can or should be made available to patients (i.e. provider psychotherapy notes, or physician intellectual property), and should remain flexible regarding such record-onboarding exceptions [5], [6].

Medical records also prove critical for research. The ONC's report emphasizes that biomedical and public health researchers "require the ability to analyze information from many sources in order to identify public health risks, develop new treatments and cures, and enable precision medicine" [4]. Though some data trickles through to researchers from clinical studies, surveys and teaching hospitals, we note a growing interest among patients, care providers and regulatory bodies to responsibly share more data, and thus enable better care for others [7], [4].

In this work, we explore a blockchain structure applied to EHRs. We build on this distributed ledger protocol originally associated with Bitcoin [8]. The blockchain uses public key cryptography to create an append-only, immutable, timestamped chain of content. Copies of the blockchain are distributed on each participating node in the network. The Proof of Work algorithm used to secure the content from tampering depends on a "trustless" model, where individual nodes must compete to solve computationally-intensive "puzzles" (hashing exercises) before the next block of content can be appended to the chain. These worker nodes are known as "miners," and the work required of miners to append blocks ensures that it is difficult to rewrite history on the blockchain.

Our MedRec blockchain implementation addresses the four major issues highlighted above: fragmented, slow access to medical data; system interoperability; patient agency; improved data quality and quantity for medical research. We build on the work of Zyskind et al. [9] to assemble references to data and encode these as hashed pointers onto a blockchain ledger. We then organize these references to explicitly create an accessible bread crumb trail for medical history, without storing raw medical data on the blockchain. Our system supplements these pointers with on-chain permissioning and data integrity logic, empowering individuals with record authenticity, auditability and data sharing. We build robust, modular APIs to integrate with existing provider databases for interoperability. A novel data-mining scheme is proposed to sustain the MedRec network and bring open, big data to medical researchers. We present MedRec not as the panacea for medical record management, but as a foray into this space to demonstrate innovative EHR solutions with blockchain technology.

2. System Implementation

2.1 Overview

For MedRec, the block content represents data ownership and viewership permissions shared by members of a private, peer-to-peer network. Blockchain technology supports the use of "smart contracts," which allow us to automate and track certain state transitions (such as a change in viewership rights, or the birth of a new record in the system). Via smart contracts on an Ethereum blockchain [10], we log patient-provider relationships that associate a medical record with viewing permissions and data retrieval instructions (essentially data pointers) for execution on external databases. We include on the blockchain a cryptographic hash of the record to ensure against tampering, thus guaranteeing data integrity. Providers can add a new record associated with a particular patient, and patients can authorize sharing of records between providers. In both cases, the party receiving new information receives an automated notification and can verify the proposed record before accepting or rejecting the data. This keeps participants informed and engaged in the evolution of their records.

MedRec prioritizes usability by also offering a designated contract which aggregates references to all of a user's patient-provider relationships, thus providing a single point of reference to check for any updates to medical history. We handle identity confirmation via public key cryptography and employ a DNS-like implementation that maps an already existing and widely accepted form of ID (e.g. name, or

social security number) to the person's Ethereum address. A syncing algorithm handles data exchange “off-chain” between a patient database and a provider database, after referencing the blockchain to confirm permissions via our database authentication server.

In the following sections we present the design principles of our distributed system and its implementation.

2.2 Blockchain Background

Originally designed for keeping a financial ledger, the blockchain paradigm can be extended to provide a generalized framework for implementing decentralized compute resources [10]. Each compute resource can be thought of as a singleton state-machine that can transition between states via cryptographically-secured transactions. When generating a new state-machine, the nodes encode logic which defines valid state transitions and upload it onto the blockchain. From there on, the blocks journal a series of valid transactions that, when incrementally executed with the state from the previous block, morph the state-machine into its current state. The Proof of Work consensus algorithm and its underlying peer-to-peer protocol secure the state-machines' state and transitioning logic from tampering, and also share this information with all nodes participating in the system. Nodes can therefore query the state-machines at any time and obtain a result which is accepted by the entire network with high certainty.

This transaction-based state-machine generalization of the blockchain is informally referred to as smart contracts. Ethereum is the first to attempt a full implementation of this idea. It builds into the blockchain a Turing-complete instruction set to allow smart-contract programming and a storage capability to accommodate on-chain state. We regard the flexibility of its programming language as an important property in the context of EHR management. This property can enable advanced functionality (multi-party arbitration, bidding, reputation, etc.) to be coded into our proposed system, adapting to comply with differences in regulation and changes in stakeholders needs.

We utilize Ethereum's smart contracts to create intelligent representations of existing medical records that are stored within individual nodes on the network. We construct the contracts to contain metadata about the record ownership, permissions and data integrity. The blockchain transactions in our system carry cryptographically signed instructions to manage these properties. The contract's state-transition functions carry out policies, enforcing data alternation only by legitimate transactions. Such policies can be designed to implement any set of rules which govern a particular medical record, as long as it can be represented computationally. For example, a policy may enforce that separate transactions representing consent are sent from both patients and care providers, before granting viewing permissions to a third party.

To navigate the potentially large amount of record representations, our system structures them on the blockchain by implementing three types of contracts. Figure 1 illustrates the contract structures and relationships.

2.3 Smart Contract Structures

2.3.1 Registrar Contract (RC)

This global contract maps participant identification strings to their Ethereum address identity (equivalent to a public key). We intentionally use strings rather than the cryptographic public key identities directly, allowing the use of already existing form of ID. Policies coded into the contract can regulate registering new identities or changing the mapping of existing ones. Identity registration can thus be restricted only to certified institutions. The RC also maps identity strings to an address on the blockchain, where a special contract described below, called the Summary Contract, can be found.

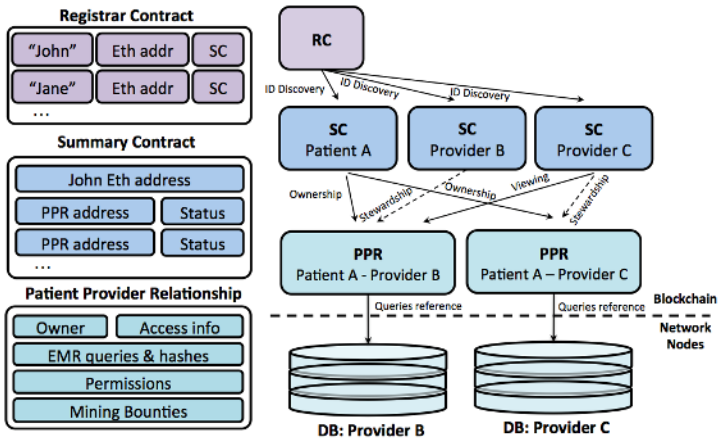

Figure 1. MedRec smart contracts on the left, showing data content for each contract type. Sample relationship graph between contracts and network nodes on the right.

2.3.2 Patient-Provider Relationship Contract (PPR)

A Patient-Provider Relationship Contract is issued between two nodes in the system when one node stores and manages medical records for the other. While we use the case of care provider and patient, this notion extends to any pairwise data stewardship interaction. The PPR defines an assortment of data pointers and associated access permissions that identify the records held by the care provider. Each pointer consists of a query string that, when executed on the provider's database, returns a subset of patient data. The query string is affixed with the hash of this data subset, to guarantee that data have not been altered at the source. Additional information indicates where the provider's database can be accessed in the network, i.e. hostname and port in a standard network topology. The data queries and their associated information are crafted by the care provider and modified when new records are added. To enable patients to share records with others, a dictionary implementation (hash table) maps viewers' addresses to a list of additional query strings. Each string can specify a portion of the patient's data to which the third party viewer is allowed access.

Our prototype demonstrates this design with SQL data queries. In a simple case, the provider references the patient's data with a simple SELECT query conditioned on the patient's address. For patients, we designed a tool which allows them to check off fields they wish to share through our graphical interface. Under the hood, our system formulates the appropriate SQL queries and uploads them to the PPR on the blockchain. Note that by using generic strings our design can robustly interface with any string queried database implementation. Hence, it can conveniently integrate with existing provider data storage infrastructure. At the same time, patients are enabled with fine-grained access control of their medical records, selecting essentially any portion of it they wish to share.

2.3.3 Summary Contract (SC)

This contract functions as a bread crumb trail for participants in the system to locate their medical record history. It holds a list of references to Patient-Provider Relationship contracts (PPRs), representing all the participant's previous and current engagements with other nodes in the system. Patients, for instance, would have their SC populated with references to all care providers they have been engaged

with. Providers, on the other hand, are likely to have references to patients they serve and third-parties with whom their patients have authorized data sharing. The SC persists in the distributed network, adding crucial backup and restore functionality. Patients can leave and rejoin the system multiple times, for arbitrary periods, and always regain access to their history by downloading the latest blockchain from the network. As long as there are nodes participating in the network, the blockchain log is maintained.

The SC also implements functionality to enable user notifications. Each relationship stores a status variable. This indicates whether the relationship is newly established, awaiting pending updates and has or has not acknowledged patient approval. Providers in our system set the relationship status in their patients' SC whenever they update records or as part of creating a new relationship. Accordingly, the patients can poll their SC and be notified whenever a new relationship is suggested or an update is available. Patients can accept, reject or delete relationships, deciding which records in their history they acknowledge.

Our prototype ensures that accepting or rejecting relationships is done only by the patients. To avoid notification spamming from malicious participants, only providers can update the status variable. These administration principles can be extended, adding additional verifications to confirm proper actor behavior.

2.4 System Node Description

We design the components of our system nodes to integrate with existing EHR infrastructure. We assume that many nodes, and in particular care providers, already trustfully manage databases with patient data stored on servers with network connectivity. Our design introduces four software components: Backend Library, Ethereum Client, Database Gatekeeper and EHR Manager. These can be executed on servers, combining to create a coherent, distributed system. We provide a prototype implementation of these components that integrates with a SQLite database and is managed through our web user interface. Notably, any provider backend and user interface implementations can participate in the system by employing the modular interoperability protocol as defined through our blockchain contracts.

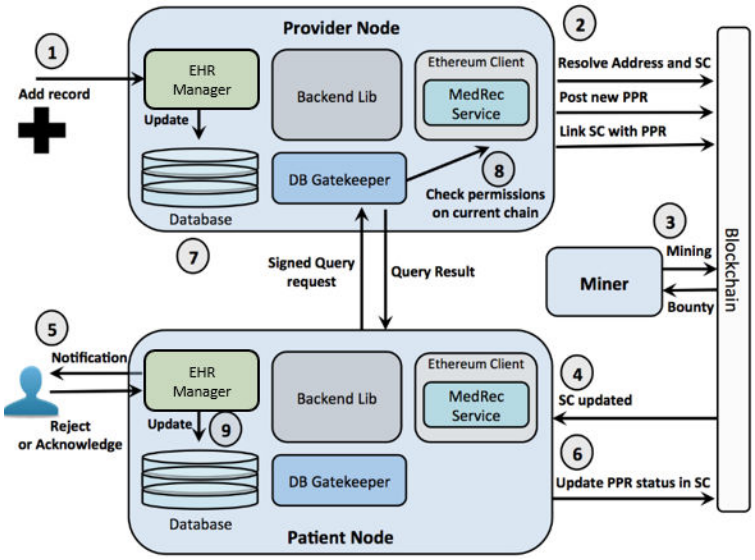

Figure 2. System orchestration example: provider adds a record for new patient.

Patient nodes in our system contain the same basic components as providers. An implementation of these can be executed on a local PC or even a mobile phone. Their local database can be one of many lightweight database implementations. The databases can function merely as cache storage of the patient's medical data. Missing data can be retrieved from the network at any time by following the node's Summary Contract.

2.5 Primary Software Modules

2.5.1 Backend API Library

We construct multiple utilities, bundled in a backend library, to facilitate the system's operation. Our library abstracts the communications with the blockchain and exports a function-call API. Record management applications and their user interfaces can thus avoid the hurdles of working directly with the blockchain. One such hurdle is verifying that each sent transaction is accepted with high confidence by the network. Our library automatically handles the uncertainty of when transactions are mined and deals with cases when they are discarded. The backend library interacts with an Ethereum client to exercise the low-level formatting and parsing of the Ethereum protocol.

Steps 1 and 2 in Figure 2 illustrate our backend implementation of a scenario where a provider adds a record for a new patient. Using the Registrar Contract on the blockchain, the patient's identifying information is first resolved to their matching Ethereum address and the corresponding Summary Contract is located. Next, the provider uploads a new PPR to the blockchain, indicating their stewardship of the data owned by the patient's Ethereum address. The provider node then crafts a query to reference this data and updates the PPR accordingly. Finally, the node sends a transaction which links the new PPR to the patient's Summary Contract, allowing the patient node to later locate it on the blockchain.

2.5.2 Ethereum Client

This component implements the full functionality required to join and participate in the Ethereum blockchain network. This handles a broad set of tasks, such as connecting to the peer-to-peer network, encoding and sending transactions and keeping a verified local copy of the blockchain. For our prototype implementation we use PyEthereum and the PyEthApp client.

We modify the client to be aware of our mapping of identity and addresses. We then implement a service to locate the node's Summary Contract (SC), via Registrar Contract address lookup. This service runs continuously within the client to monitor real-time changes to the SC. In the event of an update, the service signals the EHR Manager to issue a user notification and, if necessary, sync the local database.

Steps 4 to 6 in Figure 2 continue the use case described above from the patient node perspective. The patient's modified Ethereum client continuously monitors her SC. Once a new block is mined with the newly linked PPR, the client issues a signal which results in a user notification. The user can then acknowledge or decline her communication with the provider, updating the Summary Contract accordingly. If the communication is accepted, our prototype implementation automatically issues a query request to obtain the new medical data. It uses the information in the new PPR to locate the provider on the network and connect to its Database Gatekeeper server.

2.5.3 Database Gatekeeper

The Database Gatekeeper implements an off-chain, access interface to the node's local database, governed by permissions stored on the blockchain. The Gatekeeper runs a server listening to query requests from clients on the network. A request contains a query string, as well as a reference to the blockchain PPR that warrants permissions to run it. The request is cryptographically signed by the issuer, allowing the gatekeeper to confirm identities. Once the issuer's signature is certified, the gatekeeper checks the blockchain contracts to verify if the address issuing the request is allowed access to the query. If the address checks out, it runs the query on the node's local database and returns the result over to the client.

Steps 7 to 9 in Figure 2 illustrate how a patient retrieves personal data from the provider node. Note that our components similarly support third-parties retrieving patient-shared data: the patient selects data to share and updates the corresponding PPR with the third-party address and query string. If necessary, the patient's node can resolve the third party address using the Registrar Contract on the blockchain. Then, the patient node links their existing PPR with the care provider to the third-party's Summary Contract. The third party is automatically notified of new permissions, and can follow the link to discover all information needed for retrieval. The provider's Database Gatekeeper will permit access to such a request, corroborating that it was issued by the patient on the PPR they share.

2.5.4 EHR Manager

We tie together all the software components previously mentioned with our EHR management and user interface application. The application renders data from local SQLite databases (designed to be interchangeable with other DB software) for viewing, and presents the users with update notifications, and data sharing and retrieval options. Our user interface prioritizes intuitive, crisp, and informative design, as recommended by the Department of Veteran Affairs and ONC's Blue Button design competition [11]. Our application is conveniently accessed through a web interface, built on a python backend framework. We are especially cognizant of compatibility for mobile devices, as modern users expect easy access and high quality experiences while on-the-go.

2.6 MedRec Blockchain Mining

We incentivize "miners" to participate in the network and contribute their computational resources to achieve a trustworthy, gradual advancement of the chain. We propose a model that engages the healthcare community in network stewardship—MedRec brings medical researchers and health care stakeholders to mine in the network. In return, the network beneficiaries, i.e. providers and patients, release access to aggregate, anonymized medical data as mining rewards. We explore this idea in our prototype by implementing a special function in the PPR contract. It requires care providers to attach a bounty query to any transaction they send updating the PPR. For example, this bounty query can be formulated to return the average iron levels in blood tests done by the provider, across all patients, in the previous week. When the block containing the record-update transaction is mined, the mining function automatically appends the block's miner as the owner of the bounty query. The miner can then collect it by simply issuing a request for this bounty to the provider's Database gatekeeper. Because it is signed by the provider as part of the transaction, the bounty query is safe from malicious alterations. This "bounty query" or data reward for mining enables medical researchers to access population-level insights into medical treatment and healthcare outcomes, potentially revolutionizing how data is gathered and accessed for research purposes. We envision future updates to the mining model where miners can specify preferences for demographic cohorts and features of the data they are looking for, in order to enable precision medicine and targeted research (while still preserving the privacy of the patients).

3. Prototype Evaluation

MedRec gives patients an immutable log of their medical history, which is not only comprehensive, but also accessible and credible. This restores patient agency, as participants are now more fully informed of their medical history and any modifications to it. Through permission management on the blockchain, we enable patient-vetted data exchange between medical jurisdictions and an interoperable content management system for the physicians supervising these records. The blockchain ledger keeps an auditable history of medical interactions between patients and providers, likely relevant

for regulators and payers (e.g. insurance) in the future. Below, we consider the security, privacy and interoperability implications of this project and discuss our in-situ deployment testing.

First, on robustness and security: our blockchain implementation enjoys several key properties of decentralization. MedRec enjoys a strong failover model, relying on the many participating entities in the system to avoid a single point of failure. Medical records are stored locally in separate provider and patient databases; copies of authorization data are stored on each node in the network. Because both the raw medical data and global authorization log stay distributed, our system does not create a central target for content attack—a crucial consideration in an age of cyberattacks and data leaks. Though some blockchains experience robustness challenges from a scaling limit on the "block size" or storage capacity [12], these parameters can be modified to optimize for other performance requirements in a private blockchain network. Notably, MedRec does not claim to address the security of individual provider databases—this must still be managed properly by the local IT system admin. Nor does MedRec attempt to solve the Digital Rights Management [13] problem of undesired data copying, as our system assumes provider nodes that are bound by external regulation governing data copying in the medical use case, e.g. HIPAA.

Regarding privacy, use of blockchain technology introduces several limitations. The pseudonymous property of transactions currently allows for data forensics, or inferring patterns of treatment from frequency analysis. Without any disclosure of name or PII, one could infer that some entity has repeatedly interacted with another network entity through analysis of network traffic. Improving obfuscation while preserving auditability on the blockchain is an ongoing area of exploration. One potential solution is to make the blockchain a "permissioned" structure, where only pre-approved, white-listed nodes are allowed read access to the ledger. This would prevent rogue actors from extracting frequency-based insights from the blockchain records. Furthermore, encryption can be introduced in the off-blockchain data syncing steps to safeguard against accidental or malicious content access. While outside the scope of the initial prototype (but unarguably crucial for future development), a rigorous k-anonymity analysis [14] of privacy-preserving query construction is needed, for release of the aggregated research data to medical research "miners."

Regarding interoperability: by integrating with providers' existing data storage infrastructure, we facilitate continued use of their existing systems. We believe this will ease adoption and aid compliance with HIPAA regulations. Building on the principle of interoperability, we have designed the system with flexibility to support open standards for health data exchange—be that FHIR or other flavors of HL7 proposals in the future [15]. In addition, MedRec is source agnostic, i.e. able to receive data from any number of endpoints (physician offices, hospital servers, patient home computers, et cetera). We have developed MedRec not as a proprietary system, but as a set of open APIs to facilitate EHR review and exchange. MedRec is a layer that can be added to existing provider backends (see discussion below of integration with EPIC and Cerner systems) with minimal orchestration, thanks to the embedded logic in our Database Gatekeeper utility.

To test our system's interoperability with an in-situ provider's backend systems and data files, we have partnered with Beth Israel Deaconess Medical Center (Harvard Medical School Teaching Hospital). We are evaluating MedRec's ability to smoothly intake and parse a standard clinical document, link our Database Gatekeeper utility to the relevant Beth Israel endpoint and test an end-to-end system flow from the hospital's existing user interface for physicians through our backend and out to a sample patient node.

[End of material adapted from IEEE Conference submission "MedRec: Using Blockchain for Medical Data Access and Permission Management"]

4. MedRec in the Context of National Healthcare Priorities

As mentioned in the introduction, we do not present MedRec as a panacea nor as the only blockchain-mediated solution that would be needed to achieve our stated goals of data access, patient-empowerment, interoperability and improved medical research. In the analysis below, we refer to MedRec by name to suggest how such a project might address national healthcare priorities, likely as part of a larger suite of blockchain solutions to which we hope to contribute.

Most importantly, the MedRec model restores comprehensive patient agency over healthcare information—across providers and treatment sites, empowering citizens with the data they need to make informed decisions around their care. By giving patients a long-term, trusted log of their information with data sharing functionality built-in, the MedRec system directly addresses the ONC Interoperability Roadmap's first Outcome: "Individuals have access to longitudinal electronic health information, can contribute to the information, and can direct it to any electronic location" [16]. As envisioned by the Precision Medicine Initiative (PMI), the MedRec patient record would reflect the many facets of health data, by accepting not just physician data, but also data from the patient's Fitbit, Apple HealthKit, 23andMe profile, and more. Patients can build a holistic record of their medical data and authorize others for viewership, such as physicians providing a second opinion or family members and care guardians.

MedRec data can also feed into emerging technologies for predictive analytics, allowing patients to learn from their family histories, past care and conditions to better prepare for healthcare needs in the future. By employing open APIs like MedRec, machine learning and data analysis layers could be added to repositories of healthcare data to enable a true "learning health system" [16]. Due to the linked interoperability between provider databases in a MedRec network, better-unified access to data could facilitate a wide range of trend discovery. MedRec's modularity could support an additional analytics layer for disease surveillance and epidemiological monitoring, physician alerts if patients repeatedly fill and abuse prescription access (e.g. part of the national problem with narcotics abuse [17]), personal dashboards that show patients emerging trends in their own health, etc. In this respect, MedRec enables a service-oriented architecture (SOA) as outlined in the ONC Roadmap's "Secure, Standard Services" [16].

MedRec's community model, where medical researchers (and potentially other regulated stakeholders in the healthcare industry) can obtain insightful, population-wide data on medical treatment offers an unprecedented opportunity to achieve goals for precision medicine and evidence-based research. Such a system would facilitate the Patient-Centered Outcomes Research Institute's goals for comparative clinical effectiveness research [18], by linking the patients within a particular clinical cohort with both granular and long-term medical history, thus enabling a better understanding of patient outcomes across treatment groups and over time. By leveraging a data orchestration system like MedRec where the records would already be gathered, organized and available for analysis, this type of research can be achieved with significantly less overhead than traditional research trials, which often require expensive recruitment procedures and in-person access to patients. This ability to carry out longitudinal studies on MedRec user cohorts directly addresses both the ONC Interoperability Roadmap stated Outcomes [16] and the PMI's goal for a national research cohort [19].

The MedRec smart contract structure serves as one model for a "Health Care Directory and Resource Location," secured with public key cryptography and enabled with crucial properties of provenance and data integrity. This blockchain directory model supports the ability to "grow and change dramatically throughout its lifetime— adding new participants and changing organizational relationships" through stateful updates to the smart contracts [16]. A blockchain log could provide clarity for

communicating authorization "across the Health IT ecosystem," and an audit log for subsequent inquiries into use of such permissions and access patterns. With this functionality, the system would serve as a "Consistent Representation of Authorization to Access Electronic Health Information" [16].

Fundamentally, the MedRec project strives to enable Precision Medicine and holistic understanding of patient medical status without creating a centralized repository of data. Centrally-stored data has often proved disastrous in our modern age of cyberattacks and data leaks. Therefore, MedRec leverages a decentralized, blockchain architecture to enable local, separate storage but coordinated viewing of the data from the patient perspective. We believe MedRec fits squarely in the White House's goals for the ONC to "support the development of interoperability standards and requirements that address privacy and enable secure exchange of data across systems" [20]. Because MedRec is a system of open APIs, we hope to integrate with other key layers in the healthcare IT stack of the future.

5. Future Work

As we look to take MedRec from a research prototype to a meaningful tool for enterprise, government and patient use, we have identified several thrusts of future work. First, we continue our process of actively engaging with healthcare stakeholders across the industry, from hospitals and provider offices, to pharmaceutical companies, to insurance companies, to healthcare startups, U.S. Government institutions and more. We are currently in the process of gathering functionality requirements and additional use-case scenarios from the Department of Veterans Affairs, Kaiser Permanente, Merck & Co., Beth Israel Deaconess Medical Center and others to improve the design of all aspects of the MedRec system. In future months, we hope to complete additional rounds of security testing, including third-party penetration testing and a bug bounty program, as outlined in the ONC Roadmap's guidelines for "Ubiquitous, Secure Network Infrastructure" [16].

Though the MedRec backend is already designed to be flexible with many database architectures, we are exploring custom integration requirements for InterSystems Caché technology, which underpins many hospital backends across the nation and supports EPIC's record management platform [21]. Our goal is to make MedRec an interoperability layer that can be seamlessly added to existing EPIC, Cerner, et cetera deployments, building on the open standards development collaboration "Sync for Science" between the NIH and ONC [22].

6. Conclusion

The MedRec prototype provides a proof-of-concept system, demonstrating how principles of decentralization and blockchain architectures could contribute to secure, interoperable EHR systems. Using Ethereum smart contracts to orchestrate a content-access system across separate storage and provider sites, the MedRec authentication log governs medical record access while providing patients with comprehensive record review, care auditability and data sharing. We demonstrate an innovative approach for integrating with providers' existing systems, prioritizing open APIs and network structure transparency. We look forward to continued work on the MedRec project infrastructure, following the ONC's call for policy and technical components of an interoperable health IT stack. We remain committed to the principles of open source software and will release our research framework on GitHub as a platform for further development in the fall of 2016.

Acknowledgment

We thank the MIT Digital Currency Initiative, the MIT Media Lab Consortium and the Beth Israel Deaconess Medical Center Information Systems group for their support. We also thank and recognize MIT student Thiago Vieira for his early code contributions.

References

[1] "Who Owns Medical Records: 50 State Comparison." *Health Information and the Law*. George Washington University Hirsh Health Law and Policy Program. Aug. 20, 2015. [Online] Available: http://www.healthinfolaw.org/comparative-analysis/who-owns-medical-records-50-state-comparison

[2] U.S. Department of Health and Human Services, Office of Civil Rights. (2013). 45 CFR Parts 160, 162, and 164. "HIPAA Administrative Simplification." [Online] Available: http://www.hhs.gov/sites/default/files/hipaa-simplification-201303.pdf

[3] Mandl, Kenneth D., David Markwell, Rhona MacDonald, Peter Szolovits, and Isaac S. Kohane. "Public Standards and Patients' Control: how to keep electronic medical records accessible but private." *Bmj* 322, no. 7281 (2001): 283-287.

[4] Office of the National Coordinator for Health Information Technology. (2015). Report to Congress. "Report on Health Information Blocking." [Online] Available: https://www.healthit.gov/sites/default/files/reports/info_blocking_040915.pdf

[5] "Individuals' Right Under HIPAA to Access their Health Information 45 CFR § 164.524." U.S. Department of Health and Human Services. [Online] Available: http://www.hhs.gov/hipaa/for-professionals/privacy/guidance/access/ . Accessed: Aug. 8, 2016.

[6] Grossmann, Claudia, W. Alexander Goolsby, LeighAnn Olsen, and J. Michael McGinnis. Institute of Medicine of the National Academies. "Clinical Data as the Basic Staple of Health Learning: Creating and Protecting a Public Good." *Workshop Summary (Learning Health System Series)*. National Academies Press, (2010).

[7] Kish, Leonard J., and Eric J. Topol. "Unpatients [mdash] why patients should own their medical data." *Nature biotechnology* 33, no. 9 (2015): 921-924.

[8] Nakamoto, Satoshi. "Bitcoin: A peer-to-peer electronic cash system." (2008).

[9] Zyskind, Guy, and Oz Nathan. "Decentralizing privacy: Using blockchain to protect personal data." In *Security and Privacy Workshops (SPW),* (2015) *IEEE*, pp. 180-184.

[10] Wood, Gavin. "Ethereum: A secure decentralised generalised transaction ledger." *Ethereum Project Yellow Paper* (2014).

[11] "The Patient Record: health design challenge." The Office of the National Coordinator for Health Information Technology, U.S. Department of Veterans Affairs. Jan. 2013. [Online] Available: http://healthdesignchallenge.com/

[12] Croman, Kyle, Christian Decker, Ittay Eyal, Adem Efe Gencer, Ari Juels, Ahmed Kosba, Andrew Miller, Prateek Saxena, Elaine Shi, and Emin Gün. "On scaling decentralized blockchains." In *Proc. 3rd Workshop on Bitcoin and Blockchain Research* (2016).

[13] "Digital Rights Management and Libraries." American Library Association. [Online] Available: http://www.ala.org/advocacy/copyright/digitalrights. Accessed Aug. 4, 2016.

[14] Sweeney, Latanya. "K-anonymity: A model for protecting privacy." *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems* 10, no. 05 (2002): 557-570.

[15] "FHIR Overview." HL7 International. Oct. 2015. [Online] Available: https://www.hl7.org/fhir/overview.html

[16] Office of the National Coordinator for Health Information Technology. (2015). Version 1.0. "Connecting Health and Care for the Nation: A shared nationwide interoperability roadmap." [Online] Available: https://www.healthit.gov/sites/default/files/hie-interoperability/nationwide-interoperability-roadmap-final-version-1.0.pdf

[17] "About the Epidemic." U.S. Department of Health & Human Services. [Online] Available: http://www.hhs.gov/opioids/about-the-epidemic/. Accessed Aug. 4, 2016.

[18] "Research We Support." Patient-Centered Outcomes Research Institute (pcori). [Online] Available: http://www.pcori.org/research-results/research-we-support. Accessed Aug. 4, 2016.

[19] "Precision Medicine Initiative Cohort Program." National Institutes of Health. [Online] Available: https://www.nih.gov/precision-medicine-initiative-cohort-program. Accessed Aug. 4, 2016.

[20] "Fact Sheet: President Obama's Precision Medicine Initiative." The White House Briefing Room. Jan. 30, 2015. [Online] Available: https://www.whitehouse.gov/the-press-office/2015/01/30/fact-sheet-president-obama-s-precision-medicine-initiative.

[21] "InterSystems Unveils Major New Release of Caché." InterSystems. Feb. 25, 2015. [Online] Available: http://www.intersystems.com/who-we-are/newsroom/news-item/intersystems-unveils-major-new-release-cache/.

[22] "Fact Sheet: Obama Administration Announces Key Actions to Acclerate Precision Medicine Initiative." The White House Briefing Room. Feb. 25, 2016. [Online] Available: https://www.whitehouse.gov/the-press-office/2016/02/25/fact-sheet-obama-administration-announces-key-actions-accelerate