MyMEDIS: a new medical data storage and access system

Control Your History to Control Your Future

A blockchain-supported system to enhance or replace current methods of storage and access of electronic health records

Aram Kovach, Gabriel Ronai

February 2018

Abstract

We propose a blockchain-supported system that aspires to give control to patients over their existing medical records and health related data, while making them instantly available everywhere. The system utilizes distributed storage technology for redundancy and availability, and strong cryptographic encryption to ensure confidentiality of the content uploaded in the form of medical records, diagnostic imaging or IoT data.

Table of Contents

This document is for informational purposes only and does not constitute an offer or solicitation to sell shares or securities in MEDIS or any related or associated company. Any such offer or solicitation will be made only by means of a confidential offering memorandum and in accordance with the terms of all applicable securities and other laws.

Introduction

What is the problem with the current models, why is this system different and what are the benefits of implementing it?

Philosophical background

The Cartesian dualism

René Descartes made ontological space for modern medicine by separating body from mind – while mind is superior to body as it constitutes the uniqueness of the human soul (the province of theology), body is inferior to mind as it is mere matter. Medicine simply investigated(s) the body as machine. While Cartesian dualism dominates clinical approaches to medical research and treatment, the legitimacy of the split between mind and body has been consistently challenged from a variety of perspectives. (*Philosophy of Medicine - Wikipedia*)

Similar to this dualism in the approach to medical treatment, we can observe a dualism in the handling and accessibility of data in current medical record keeping systems. While the originator of the data in question is the patient, he/she has no control over how, where that data is kept, and how it is used or accessed after being created.

In this document we propose a new system to store, manage and access electronic Personal Health Information (e-PHI). Participation is voluntary for patients, healthcare providers and other entities who interact with the system. Incentives and advantages are offered to end users and institutions.

What is Blockchain technology?

A blockchain – originally block chain – is a distributed database that is used to maintain a continuously growing list of records, called blocks. Each block contains a timestamp and a link to a previous block. A blockchain is typically managed by a peer-to-peer network collectively adhering to a protocol for validating new blocks. By design, blockchains are inherently resistant to modification of the data. Once recorded, the data in any given block cannot be altered retroactively without the alteration of all subsequent blocks and the collusion of the network. Functionally, a blockchain can serve as "an open, distributed ledger that can record transactions between two parties efficiently and in a verifiable and permanent way. The ledger itself can also be programmed to trigger transactions automatically." (*Blockchain - Wikipedia*)

What is Hyperledger?

Hyperledger is an open source collaborative effort created to advance cross-industry blockchain technologies. It is a global collaboration, hosted by The Linux Foundation, including leaders in finance, banking, Internet of Things, supply chains, manufacturing and Technology.

A blockchain is a peer-to-peer distributed ledger forged by consensus, combined with a system for “smart contracts” and other assistive technologies. Together these can be used to build a new generation of transactional applications that establishes trust, accountability and transparency at their core, while streamlining business processes and legal constraints.

Think of it as an operating system for marketplaces, data-sharing networks, micro-currencies, and decentralized digital communities. It has the potential to vastly reduce the cost and complexity of getting things done in the real world. (“About - Hyperledger”)

What is chaincode?

Chaincode is specific to Hyperledger and different from the Smart Contract programming language used by other blockchain solutions, such as the Solidity language and compiler created by the Ethereum Foundation.

Chaincode is software defining an asset or assets, and the transaction instructions for modifying the asset(s). In other words, it’s the business logic. Chaincode enforces the rules for reading or altering key value pairs or other state database information. Chaincode functions execute against the ledger’s current state database and are initiated through a transaction proposal. Chaincode execution results in a set of key value writes (write set) that can be submitted to the network and applied to the ledger on all peers. (*Hyperledger Fabric Model — Hyperledger-Fabricdocs Master Documentation*)

What is a token?

Ethereum tokens are simply digital assets that are being built on top of the Ethereum blockchain. They benefit from Ethereum’s existing infrastructure instead of developers having to build an entirely new blockchain. They also strengthen the Ethereum ecosystem by driving demand for ether, the native currency of Ethereum, needed to power the smart contracts.

Ethereum tokens can represent anything from a physical object like gold to a native currency used to pay transaction fees. In the future, tokens may even be used to represent financial

instruments like stocks and bonds. The properties and functions of each token are entirely subject to its intended use. Tokens can have a fixed supply, constant inflation rate, or even a supply determined by a sophisticated monetary policy. Tokens can be used for a variety of purposes such as paying to access a network or for decentralized governance over an organization. (Xie)

MyMEDIS has created the MediCoin token on the public Ethereum network with the intent to facilitate reward and transfers of value (payments) between participants, independent from the isolated Hyperledger and storage network.

Preliminaries and problem statement

Current infrastructure

Currently, healthcare institutions own and maintain distinct EMR systems that are not communicating with each other, or do so at a very rudimentary level. Over the course of their lives patients visit multiple institutions, thus their records are spread across multiple disconnected systems, stored by each separate institution and practically inaccessible to the patient.

Additionally, access to these EMR entries is controlled by the institutions using their chosen systems on an IT infrastructure that does not allow external access. When switching healthcare providers, patients or doctors must request copies of their records to be transferred from one institution to the other, often by means of traditional post. In case of emergency, the responder does not know of any preexisting conditions or treatments that may affect their procedures.

Current patient-institution relationships

Traditionally, healthcare providers prefer not not to give any access to digital EMR records to the patients. Paper copies of care documentation are usually provided after visits, in post-operative discharge paperwork, etc. Understandable concerns regarding forgery of records resulted in total prohibition of digital access by the patient to electronic healthcare records.

Any ability to alter records from the patient side renders the stored data unreliable in regard to any further use in treatment or research. Additionally, it poses great risk from legal and healthcare perspectives. Therefore we believe that patients should have access to their records, but only in a read-only manner. Patients should also have the right to allow or block access to their records for other participants of the healthcare industry. Furthermore, if the EMR has research value, the patient should be reimbursed accordingly when their anonymized records are used by other entities for any purpose.

HIPAA Regulations and Compliance

In order to protect the privacy and security of certain health information, the U.S. Department of Health and Human Services (HHS) published what are commonly known as the HIPAA Privacy Rule and the HIPAA Security Rule. The Privacy Rule, or Standards for Privacy of Individually Identifiable Health Information, establishes national standards for the protection of certain health information. The Security Standards for the Protection of Electronic Protected Health Information (the Security Rule) establish a national set of security standards for protecting certain health information that is held or transferred in electronic form. The Security Rule operationalizes the protections contained in the Privacy Rule by addressing the technical and non-technical safeguards that organizations called “covered entities” must put in place to secure individuals’ “electronic protected health information” (e-PHI). Within HHS, the Office for Civil Rights (OCR) has responsibility for enforcing the Privacy and Security Rules with voluntary compliance activities and civil money penalties.

The HIPAA Privacy Rule protects the privacy of individually identifiable health information, called protected health information (PHI), as explained in the Privacy Rule. The Security Rule protects a subset of information covered by the Privacy Rule, which is all individually identifiable health information a covered entity creates, receives, maintains or transmits in electronic form. The Security Rule calls this information “electronic protected health information” (e-PHI).

Specifically, covered entities must:

- Ensure the confidentiality, integrity, and availability of all e-PHI they create, receive, maintain or transmit;
- Identify and protect against reasonably anticipated threats to the security or integrity of the information;
- Protect against reasonably anticipated, impermissible uses or disclosures; and
- Ensure compliance by their workforce.

The Rules and Guidelines further discuss Risk Analysis and Management, Administrative Safeguards, Technical Safeguards and various Policies, Procedures and Documentation Requirements that covered entities are required to comply with.

HIPAA defines a Business Associate as anyone who has access to patient information, whether directly, indirectly, physically or virtually. This includes any organization that provides support in the treatment, payment or operations associated with protected health information.

Business associates include:

IT providers, health applications

Telephone service provider, document management and destruction

Accountant, lawyer or other service provider

Business Associates have the responsibility to achieve and maintain HIPAA compliance in terms of all of the internal, administrative, and technical safeguards.

Business Associates must sign a Business Associate Agreement with a covered entity (such as a doctor) or other business associate (such as an application developer) to document the relationship and agree upon the terms as outlined in the requirements in HIPAA.(Office For Civil, “Summary of the HIPAA Security Rule”)

System overview

Hereby we propose a different approach to the storage, access and monetization of medical data, where EMR entries are kept in a distributed storage layer in an encrypted form. The system does not replace existing EMR solutions, but it offers tremendous value to the patient by providing centralized storage of their medical records and control over who and how gets permission to access these records in this system.

Implementation goals

From the user's perspective, the system has to be intuitive and easy to use in order to avoid frustration. By registering and using the system, patients get access to a copy of their records regardless of where those records were created.

From the operator's perspective, the MyMEDIS system must provide sufficient storage space for all uploaded data, apply proper security measures, implement activity and audit trail logging as well as offer accessible customer support tools. For researchers, the system must be able to provide valuable anonymized data sets of highest quality and of the largest size possible. For corporate partners, the system must provide proper tools for managing advertising campaigns, marketing tools for reaching out to customers as incentives for participation.

Functional overview

On the highest level, the proposed solution has to satisfy a set of simple requirements:

(1) ability to acquire (by proactively collecting or passively accepting) new EMR data

(2) provide reliable access to said data based on preset access control settings

(3) securely store the data while offering high level of availability

(4) offer additional benefits and added value to effectively compete with existing EMR storage and management systems.

System implementation

Neo4j & Ethereum vs Hyperledger & Couchbase datastore

We have decided to replace the blockchain platform and migrate the application logic to Hyperledger to utilize the built-in access control, modeling, transactional and querying abilities of the Hyperledger Fabric architecture.

MyMEDIS will use a Couchbase Server cluster to store the encrypted records and metadata instead of the Swarm and Neo4j approach described in the previous iteration of system design.

Implementation overview

The system will have several distinct subsystems that communicate with each other over encrypted channels: Datastore, Blockchain, Backend Application and End User Applications

Datastore

A data storage solution capable of storing large amount of files and can be selected based on HIPAA compliance, location, jurisdiction, security features etc. We have evaluated and selected Couchbase Server based on its local clustering and replication design, mobile synchronization and cross-datacenter replication capabilities. The storage subsystem can be re-evaluated and

Blockchain

The deployed Hyperledger network running on nodes hosted by MyMEDIS and contributor companies.

Backend Application

The application that communicates with the Hyperledger REST API, the data store interface, exposes the API endpoints and provides backend functionality for communication with the End User Application.

End User Applications

The collection of native applications (Android, iOS, Windows Mobile) used by the patients to access their records, grant permissions, read notifications, post reviews, check wallet balance, manage offers, invitations and discount coupons etc. Also provides authentication, local storage and synchronization functionality.

Web interface

A web based interface available to patients and other participants, used to interact with various system features.

Datastore architecture

Records and files are stored in a Couchbase Server cluster operated by MyMEDIS. This approach ensures a reliable service and has several properties that make it ideal for this project. Document and key-value data models make it easy to store records in JSON format, either directly mapped from FHIR resources or converted from other formats like XML. Couchbase also supports attachments, a feature we will use to hold diagnostic imaging and other binary files while still making them easy to find and retrieve.

We attempt to build a secondary database cluster from nodes running Couchbase instances in Docker containers at contributing partners. These nodes connect over a VPN tunnel and create a virtual local network that will host a replica of the database using built-in Couchbase Cross-Datacenter Replication tools. This geographically distributed cluster poses additional challenges regarding network connectivity and latency, but can be used as a disaster recovery solution if proven viable from a technical standpoint.

Couchbase Mobile provides NoSQL storage for mobile applications, as well as synchronization abilities with the main Couchbase Server. By using Couchbase Mobile in the patient native applications on iOS and Android, we can use the local storage space on the device to hold a copy of their own data and make it always accessible for viewing, even if their device does not have a network connection. With the current availability of public WiFi networks and unlimited carrier data plans, the patient's smartphone can synchronize and download any new data when online. Thus, the patient's device acts as a backup of his/her own medical history. (*Couchbase Mobile*)

Modeling data, permissions and logic in Hyperledger

The Hyperledger Fabric Model allows developers to define the basic elements of the network. The following components will be created and maintained by our developers:

Assets: MyMEDIS handles all EMR files and documents as assets. Ownership and permissions are defined in the network data model and ACL files.

Participants: Any entity that interacts with the system must be registered and assigned proper credentials and roles. Examples: Patient, Practitioner, Researcher, Pharma employee.

Access control, security and membership services: using the built-in ownership and ACL settings, participants can control who gains access to the assets they own.

Chaincode: business logic describing how and when transactions are executed.

At time of registration, each entity (patient, healthcare provider, institution, lab, operator) will have an account and a Participant record created in the Hyperledger network. The account is locked with the participant's passphrase entered during registration.

Each Patient model will hold a list of Participants they have given permanent write access to, as well as a list of records and participant identifiers that have read access to the owner's assets. These lists act as a trusted partner ledger for every Participant.

During storage, the system will anonymize (de-identify) the records while keeping useful information indexed in order to support meaningful data set extraction.

Data encryption

In order to maintain participant privacy and safety of the records, all EMR data kept in the storage subsystem is encrypted with a symmetric encryption key during creation. Anonymized information is also stored in order to maintain application logic and allow for lookup and indexing of records.

Each medical record asset in Hyperledger holds a pointer to the document in the datastore and the cryptographic encryption key to access the data. The symmetric key is encrypted with the patient's public key. When permission is granted to another participant to access the record, a copy of the record symmetric key is encrypted with the other participant's public key and stored in the record's model in Hyperledger. When the record is requested by any authorized participant, the key is retrieved through the Hyperledger REST API by the Dispatcher/Backend Application, decrypted with the recipient's private key, then the data is decrypted and delivered to the requesting application over secure transport channels.

Modeling resources in Hyperledger

Resource models and relationships are described in Hyperledger's proprietary CTO language. Some examples are shown below. Please note that these concepts are subject to change depending on legal limitations, local regulations, system design and safety considerations.

```
enum ParticipantType {
  o PERSON
  o ORGANIZATION
}

abstract concept Address {
  o String street default = "1200 Sample Street"
  o String city default = "New York"
  o String state default = "NY"
  o String postalCode default = "00000"
  o String country default = "US"
}
```

Simple enumerated and abstract resource models

```
abstract participant MyMedisParticipant identified by participantId {
  o String participantId
}

participant Patient extends MyMedisParticipant {
  o ParticipantType type default = "PERSON"
  // array of participants who can add records to this patient's library
  --> MyMedisParticipant[] assetUploadGranted optional
}

participant Practitioner extends MyMedisParticipant {
  o ParticipantType type default = "PERSON"
}

participant Organization extends MyMedisParticipant {
  o ParticipantType type default = "ORGANIZATION"
}
```

Participant model structure

```
asset participantProfile identified by assetId {
  o String assetId
  --> MyMedisParticipant owner
  o String firstName
  o String lastName
  o String organizationName optional
  o String emailAddress
  o String homePhoneNumber
  o String cellPhoneNumber
  o Address address
}

asset ICEProfile identified by assetId {
  o String assetId
  --> MyMedisParticipant owner
  o String NfcIdentifier optional
  o String KnownAllergies optional
  o String KnownIllnesses optional
  o Boolean isOrganDonor default = false
  o String insurancePolicyCompany optional
  o String insurancePolicyNumber optional
  o String contactPhone1
  o String contactPhone2 optional
  o String contactPhone3 optional
}
```

Asset models representing personal and emergency profiles

```
abstract concept AssetPermission {
  --> MyMedisParticipant participant
  o String decryptKey
  o String permissionLevel default = "READ"
  o DateTime grantedAt
  o DateTime expiresAt
}

asset medicalRecord identified by assetId {
  o String assetId
  o String storageIdentifier
  o String dataEncryptKey
  o String description optional
  --> MyMedisParticipant owner
  --> MyMedisParticipant createdBy
  o String source
  o DateTime createdAt
  o AssetPermission[] permissions
}
```

A simple medical record with the associated permission array model concept

The model definitions above allow for flexible scripting and setup access rules and business logic behavior.

Controlling permissions

When access is granted to a participant to read a specific record, the document key is encrypted with the recipient's public key and added to the asset's permission array in Hyperledger. The entity attempting to access a record can now decrypt the key and therefore access the document using their own private key.

When access is revoked, the permission entry is simply removed from the asset, and the participant can no longer access the decryption key and access the encrypted record.

Modeling e-PHI record data for storage

Storing the records in Couchbase as JSON documents allows us to run rich queries with low latency when serving requests coming from the application layer. A data record entry skeleton of a sample laboratory diagnostic report received from a lab participant in FHIR format could be similar to the one outlined below. The de-identified structure allows for filtering and search while keeping personal data hidden, while a data hash ensures record integrity.

```
{
    "id": "entry global identifier",
    "owner": "Participant identifier",
    "classification": {
        "code": "A",
        "display": "Class A record",
        "owner_generated": false
    },
    "date_created": "2004-01-20T05:44:14+00:00",
    "date_uploaded": "2004-02-01T15:19:21+00:00",
    "originator": "Participant identifier (organization)",
    "uploader": "Participant identifier (employee)",
    "metadata": {
        "resourceType": "DiagnosticReport",
        "status": "final",
        "active": true,
        "id": "sample.report.identifier.001",
        "description": "Sample laboratory diagnostic report for *** ***",
        "format": {
            "display": "FHIR Release 3",
            "version": "STU; v3.0.1-11917",
        },
        "managingOrganization": {
            "reference": "Organization/x001",
            "display": "Sample University Medical Centre"
        },
        "patient": {
            "age": 52,
            "residence": {
                "country_code": "US",
                "state": "Ohio"
            }
        },
        "contenthash": {
            "algo": "SHA256",
            "value": "a3b0f122d98ef3d7727dbdbdcb8ff0e3d666f.."
        },
    },
    "content": "...encrypted data here..."
}
```

Acquiring the data

Data classification

When a new data file is uploaded to the system, it will be assigned to a content class based on its origin. Multiple originator classes will exist in the system: care provider, insurer, patient,

research facility… The classification of content will be saved to the storage database along with the unique address, ownership relations and other meta-information about the data. Knowing a record's class becomes important for care providers when accessing historical entries, as data uploaded by the patient is usually not accepted as objective and thus deemed unreliable. Physicians who gain access to a patient's records will make the decision to use or discard the information in their current or future treatments.

Data originating from healthcare providers

The HIPAA guidelines do not specifically require local storage, only strict control over access to the data. These records can be viewed as reliable as the creator and uploader is an institution entity and originate from laboratories, other physicians, specialists or hospitals.

Data originating from the patient

The patient application will have the ability to import and collect additional health-related data, such as heart rate logs, daily step counts, sleep tracking data from fitness trackers, meal ingredient data and quantities for calorie-tracking, and more. Features can be implemented gradually as the system matures. As previously discussed, the information provided by the patient cannot be deemed reliable and is not intended to be used for healthcare services or life saving operations, rather than monetization and user convenience.

Collecting and importing historical records for patients

MEDIS will sign interoperability agreements with affected healthcare institutions. Patients can then ask these institutions to retrieve EMR data from their storage and upload them in electronic format to the MEDIS system or provide them in electronic format to the patient. Alternatively, after signing proper authorization and NDA documents, MEDIS personnel will request and import records on behalf of patients for a specified service fee. These historical entries will be then uploaded to the system and become a part of the patient's personal record library.

Automated collection of future records

The system will provide properly documented API endpoints for healthcare provider institutions to implement into their existing IT infrastructure and EMR storage systems. In the future, all newly created EMR records and documents should be automatically uploaded to MEDIS for the patient in a standardized format (HL7, FHIR, Epic, CareConnect) over an encrypted connection using SSL/HTTPS protocol.

The MyMEDIS application will also expose standard FHIR endpoints for interaction with external data storage and management applications. The endpoints will be accessible with an API key authentication system. (*Http - FHIR v3.0.1*)

Storing the data

How much data are we talking about?

How big can an EMR become?	The average size of an electronic health record, not counting images is between 1 and

• Less than 1MB for a relatively healthy adult patient • 1MB per page for scanned, paper-based data in TIFF format • 40MB without images for a patient with major medical issues • Approximately 300MB per image if picture archiving and communication system (PACS) images are included • 3GB minimum (no annotations) per patient if genome data is included Source: ("Storage Gets a Dose of Medical Data - Storage Technology Magazine")	40 MB. The absolute top end of the range is 3-5 GB "for a person with several health issues including images" (2011) (Halamka and Profile)

Ideally, all EMR data is modeled, stored and transferred in JSON (key-value) format to optimize and expedite network communication. Binary data such as diagnostic images, video recordings, are attached to the key-value storage as attachments and can be retrieved using methods similar to JSON-modeled EMR entries.

If records submitted by third party participants are not modeled in standard JSON (for example FHIR), additional mapping and conversion is required prior to submitting into the storage layer for permanent storage (database).

MyMEDIS will store the information related to the business logic and ledger in Hyperledger's own storage database (LevelDB or CouchDB), which is automatically replicated on all participant nodes and managed internally by the Hyperledger Fabric framework. This data does not contain medical records, it is primarily account and asset information, access control information, permissions and audit trail.

If the implementation requires localized private storage (i.e. due to geographic or legal limitations), the open but permissioned network can be replaced with a local, closed implementation of both the Hyperledger and database systems. The storage capacity and redundancy of such private network depends on the number of connected nodes and the cumulative storage available on the nodes, as well as the configuration of the underlying database cluster and hardware used.

For expanding the MyMEDIS network we propose a system comprised of off the shelf NAS (Network Attached Storage) hardware units that use a containerization solution (Docker) that will contain and run the software used by MyMEDIS. If required, the built in VPN features of the NAS appliances will be used to create a private network for communicating with the rest of the nodes. The centralized nature of such an approach is susceptible to various attacks like DDoS. Participant corporations (healthcare institutions, insurance and research companies) can also operate a node on their internal infrastructure, running either dedicated hardware or VPS (Virtual

Private Server) that conforms to the node requirements. The storage and application/ledger containers can be individually managed by MyMEDIS and software updates will be provided independently for each component.

Protecting the data

Apart from medical records, the transaction history and business logic information have to be kept safe and properly protected from unauthorized access.

Any record added to the system will be encrypted before being uploaded to the storage subsystem. The built-in hardware- or software-based volume encryption is enabled on the NAS devices before the containers are installed. The encrypted hard drive is transparent to the application and only adds a small overhead in exchange for data safety advantages. Modern hardware encryption engines simplify the process and improve response time.

Advantages of NAS volume encryption on remote nodes

- if a single hard drive is stolen, the data cannot be accessed due to the RAID configuration and volume encryption
- if the entire device is stolen, the attacker needs an account and password to access the OS on the device (typically Linux based, accessible over an SSH connection)
- if the attacker gains access to the underlying operating system and can access the database storage, he/she will still need the encryption key for each entry in order to decrypt the content in order to see the personally identifiable information

A permissioned network for security, auditing and access control

We have decided to replace the original Ethereum network design with the Hyperledger Fabric architecture due to the obvious advantages offered by the permissioned approach used in Hyperledger. With this new solution, nodes are assigned a unique cryptographic key by the network authority (operator) and as such a node is only able to connect to the network if the key is valid.

Additionally, by implementing Hyperledger's permissioned ACL (Access Control List) approach, only authorized entities will have access to the EMR entries. Access to records and files in the system is controlled and logged using a set of access rules built into the business logic.

Sign-in attempts, EMR data queries, file upload and download requests, cryptographic key requests and all other data access activity logs will be made available to any participant who requests to see an audit trail of their own data or activity. All audit trail and activity logs are broadcast as Events and thus saved as immutable records to the ledger (blockchain).

Preventing accidental data deletion

The system will implement a signed confirmation approach. Whenever a delete request for an EMR entry is requested from the system, the owner has to sign and confirm a transaction.

Physical protection

Depending on concrete implementation, protection is needed against unauthorized physical access to the servers or network appliances where the application or database are hosted. Personnel in charge at the data center facility is responsible for enforcing physical access control to the premises, or to equipment within the facility such as a locked server cage in a co-location building.

Protecting the application, ledger and data

Given the theoretically secure design of the Hyperledger subsystem and depending on the concrete implementation, MyMEDIS may allow nodes operated by third parties to connect to the network in order to increase redundancy, combined storage capacity or computation resources contributed by the participants. These participants can be rewarded in tokens based on the amount and type of resources contributed. However, by allowing third party nodes to connect to the network the system becomes open to malicious attacks, bringing the need to address these events and take preventative steps.

Since the MEDIS blockchain is an isolated and strictly controlled instance of Hyperledger, an attacker cannot perform a 51% attack against the network, as they can't connect to it in the first place. An attack conducted against the network aiming to slow down response time or completely block communication would not result in any economic gains for the attacker. In order to completely hinder the operation of the Hyperledger network, an attacker would have to conduct a Distributed Denial-of-Service (DDoS) attack against all nodes concurrently.

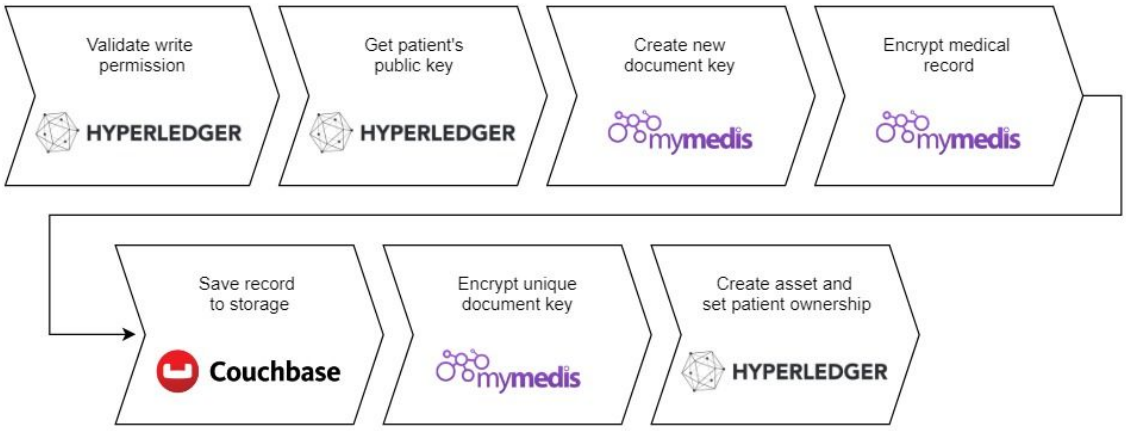
Third party storage contributors will not be able to access the content of any files, as all personally identifiable data is encrypted before being uploaded. Encryption keys are not published to any participant nor stored in Ethereum contract storage or other publicly accessible location or form.

Adding new records

When an entity submits a new request to add an EMR entry to a patient's library, the Dispatcher will first query the Hyperledger application and request confirmation that the recipient appears in the trusted partner list, thus has given this uploader permission to add records. In this

context, ‘write permission’ simply means that the uploader Participant can create new Assets where the owner is the recipient.

When a new record is added to a patient’s library, the following process takes place.



If the Hyperledger API confirms write permission, the dispatcher encrypts the uploaded file with a newly generated key and uploads it to the storage system, then saves the storage identifier, metadata, ownership and additional details as a new Asset in the Hyperledger. The patient becomes the owner of the Asset and a notification is sent to the patient’s account.

If the Hyperledger response does not confirm write permission, the record is encrypted with a newly generated key and temporarily stored in a dedicated transition area. A notification is sent to the recipient about the new request including the requestor’s identity and information about the record being uploaded.

The recipient has to decide whether to accept the request or reject it. If the request is accepted, the Dispatcher initiates a new Hyperledger transaction to update the Participant data and save the new asset ownership and the relationship information is also saved using the uploader’s participant identifier used by Hyperledger.

If the permission request is rejected, the file is removed from temporary storage and both the uploader and recipient are notified about the negative outcome.

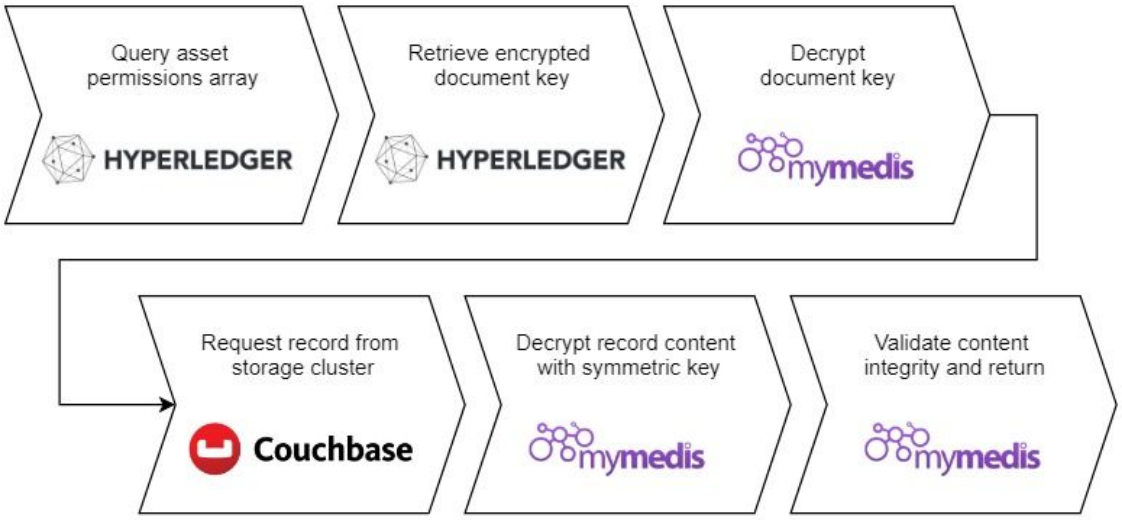
Requesting access to read a record

When an Participant requests access to an EMR entry, the Dispatcher will instruct the Hyperledger to check if the requestor has access to the record by reading the Asset permission

array. Internally, the Ledger chaincode will execute the correct function, emit logs and return the requested values through the application's Javascript API.

If the permission check result is affirmative, the Dispatcher will request the record from the database using the record's unique global identifier, decrypt it with the encryption key received from the Hyperledger and send it to the requestor for viewing.

When a record is requested by any participant other than the owner, the following process takes place.



Giving access to a record is achieved by giving access to the encryption key for that record.

Indexing the data

The system will keep information about all uploaded health records and files in a NoSQL database hosted on a cluster operated by MEDIS. The metadata stored about EMR entries covers various anonymized attributes about the record in question including ownership, size, integrity checksum, record classification, the uploader's identity and additional information needed for effective indexing and lookup.

Creating value: building a sustainable ecosystem, monetization, incentives, rewards, advertising

Monetizing records

Patients can opt in to monetize their data and set a price in tokens for their record entries. Entities interested in research data (researchers, insurance companies, universities) can request data sets from MyMEDIS containing anonymized records. MEDIS will conduct the query, data collection and export to the required format, negotiation and secure transfer of the data records from the patient to the requestor entity. By setting and publishing a token price for a specific record or a fixed price valid for all their entries, the owner (patient) agrees to give non-exclusive, revocable, read-only permission to an anonymized version of the record to all entities interacting with the system who chose to request the data contained in the record.

Data de-identification will be performed according to HHS HIPAA De-identification rules and guidance documents. (Office For Civil, “Methods for De-Identification of PHI”)

Targeted advertising

Interested entities (healthcare facilities, pharmaceutical companies, cosmetic companies) will be offered targeted advertising opportunities to patients who had a specific illness, or targeting geographic areas, age groups, genders or any combination of these. Payment for such advertising campaigns will be done in tokens to increase interest and token circulation.

Targeted clinical study participation

Research / pharmaceutical companies could search for patients matching specific age, gender, location and medical history and approach them with new clinical study opportunities. Participants will get rewarded in tokens.

Rewarding content creation

Patients who received treatment at a specific institution can be later contacted using methods built into the end user application or by traditional means (for example phone or email, if permission was granted) and asked to submit reviews and comments about the quality of service or treatment. Patients will be rewarded in tokens for adding valuable information accessible for other patients during their decision making.

Paying for treatment

When the ecosystem token reaches critical acceptance and becomes a household name, participant healthcare providers, pharmaceutical companies can set prices in tokens for their products, services or treatments. Patients will have the ability pay for treatment in tokens, should their balance hold enough to cover the cost of treatment or the price of the product offered. Institutions can convert tokens into currency on exchanges or vice-versa, generating a constant token supply for the ecosystem.

Rewarding participants

Institutions and individuals who operate a hardware or software version of the third party node will be rewarded in tokens for offering hard drive space for storage and CPU capacity for Ethereum smart contract execution.

Introduction on cryptocurrency exchanges

The MEDIS Token will be introduced on multiple exchanges in order to increase circulation and satisfy market demand.

Hardware components

In light of the requirements listed above, the system will use multiple hardware components depending on implementation requirements.

Database cluster node: a set of virtual or hardware (shared or dedicated) server device running the database container. These servers together comprise the data storage layer of the MyMEDIS, using real-time backup and replication.

Hyperledger node: a virtual or physical system that is able to run the Hyperledger container. In this context, the node will participate in the execution of the deployed application logic, hold the transaction ledger, perform any other computation tasks on the network and ensure storage of the blockchain data: ledger and chain state.

Note: it is recommended to select a NAS appliance where hardware limits permit both the Hyperledger and storage

Patient (end user) device: smartphone, tablet running iOS or Android operating system and having the End User Application installed.

Node device: a commercially available NAS device with pre-installed application components that can be purchased and operated by registered institutions interested in operating an off-site node (network participants). The device will connect to a preconfigured VPN network prior to interacting with the storage and/or Hyperledger networks. All built-in security features (hard drive encryption, data protection with RAID) are enabled before shipping.

MyMEDIS may decide to make the application software solutions available using an open source model to encourage improvements in design and security.

Looking ahead

Long term

As a long term vision, the system can be extended to provide an extended health resource locator service to healthcare systems and providers. Such a service would for example allow practitioners to perform lookups and find out whether a specific patient’s data is hosted and available in the system and act accordingly. The reverse of the process can also be

implemented, allowing patients to find providers offering specific services (dental, cosmetic surgery, etc).

Similar systems are being designed and developed on regional and national level in several countries, for example the National Record Locator Service in the United Kingdom.

> *NRLS will empower professionals, patients and communities, strengthen primary, secondary and acute care and introduce system efficiencies. This new national capability will aim to complement local digital initiatives (including Shared Care Records) and enhance the level of digital maturity across Health and Care. [...] NHS Digital intends for the NRLS to be accessible to all localities in 2018, opening up the possibility of achieving the 'holy grail of interoperability' in the months and years that follow.* (*Introduction to NRLS FHIR® API*)

Generating revenue

MyMEDIS revenue streams can be established from the following sources:

- advertising revenue from healthcare, health tourism, pharmaceutical, cosmetic and other industries
- commission from sales of anonymized data sets for statistical analysis and research
- subscription-based hosting of teleconsultation and appointment interfaces for practitioners
- transaction fees from value transfers within the ecosystem (token payments, product purchases etc.)
- insurance integration with risk mitigation
- monetizing referrals by implementing a patient review and recommendation system
- sale of MediCoin tokens on exchange
- Capital investment

Development roadmap

Phase 1: Proof of Concept – 6 months

MEDIS will lay the foundations of the system and begin development of the underlying technologies, such as a Hyperledger permissioned node network, database cluster, data model design, API and user interface specification.

The team will proceed to develop a proof of concept application and API environment to demonstrate the viability of the proposed approach.

Phase 2: Core Development and Pilot – 6 months to 12 months

Building onto the strong foundation of the previous phase, MEDIS will continue work on the system and improve features based on advisory and industry partner feedback.

Phase 3: Public release – 12 months to 18 months

MEDIS will release the first version of the end user application that allows patients to interact with the system. By this time, the main storage and backend access components are operational.

Phase 4: Monetization and industry acceptance

While continuing development, MEDIS will sign interoperability agreements with healthcare, insurance, marketing industry companies and sign them up as MEDIS ecosystem participants.

Use cases and example scenarios

Scenario 1: Patient looking for additional information about their recently diagnosed illness

A registered patient wishes to find more information, for example lab results of other patients in order to compare the values and gain more insight on the progress of their illness. Using the client program installed on their device or the web interface, the patient performs a search for the HL7 V2.x code they see on their own EMR record. From the result list, based on the record owner's visible information (age, location, gender), they decide to purchase and download one of the matching records for further viewing. After transferring the required token amount, the anonymized record is downloaded and displayed by their client program. The system will credit the record owner's balance with the correct token amount and save audit logs of the entire process.

Scenario 2: Research laboratory looking for representative data on pain medication and dosage administered during knee surgery on female patients between 40 and 60 years old

A research laboratory needs a data set containing painkiller dosage for a new research paper. The person with proper access permission on the MEDIS network will sign in using the web interface, perform a search / query of the database with the correct filters and parameters to match the records needed. Upon reviewing the list of record matches, the user selects the required entries and using a standard checkout procedure, they purchase all record data. The system will distribute the collected token revenue among the owners and make all records available for download.

Scenario 3: Pharmaceutical company looking for participants in clinical trial for new medication

For example, if a pharmaceutical company wanted to perform a clinical study, medis coins could be provided to the individuals as compensation for their participation. Up to now since medical data has been geographically dispersed and no query able fashion globally. It has been difficult to perform research where is the accuracy of the efficacy of a particular medication didn't have Standard deviation error greater than ± 5 percent, of the new drugs what would be criticized for lack of evidential proof there actually useful in curing or treating a disease. By being able to engage millions of people in a global clinical trial voluntarily, the Gaussian bell

curve typically associated with such studies would-be several times more accurate as a one standard deviation error could refocus the accuracy and efficacy of medication, because the volume of participants by scientific standards would make the results more accurate.

Scenario 4: Sharing home remedies and traditional remedy recipes

Another advantage of the medicine network would be realized in its ability to share private remedies I have worked for individuals with similar ailments. For example, an individual could trade, and Methodist coins his information regarding treatments for various Lyme disease related elements. Thereby reducing the need mine only on physician prescribed antibiotics of ever-increasing potency whose long-term use severely Impact's kidney function. There bye anybody with a similar debilitating lyme disease related immobility could volunteer and all information to others by making it available for sale if the actual remedy yields positive results.

References

"About - Hyperledger." *Hyperledger*, https://hyperledger.org/about. Accessed 20 Feb. 2018.

Blockchain - Wikipedia. http://en.wikipedia.org/w/index.php?title=Blockchain&oldid=784977253. Accessed 20 Feb. 2018.

Couchbase Mobile. https://developer.couchbase.com/mobile. Accessed 21 Feb. 2018.

Halamka, John, and View my Complete Profile. *The Cost of Storing Patient Records*. http://geekdoctor.blogspot.com/2011/04/cost-of-storing-patient-records.html. Accessed 20 Feb. 2018.

Http - FHIR v3.0.1. https://www.hl7.org/fhir/http.html. Accessed 21 Feb. 2018.

Hyperledger Fabric Model — Hyperledger-Fabricdocs Master Documentation. http://hyperledger-fabric.readthedocs.io/en/release/fabric_model.html#chaincode. Accessed 22 Feb. 2018.

Introduction to NRLS FHIR® API. https://nhsconnect.github.io/CareConnectAPI/. Accessed 21 Feb. 2018.

Office For Civil. "Methods for De-Identification of PHI." *HHS.gov*, US Department of Health and Human Services, 6 Nov. 2015, https://www.hhs.gov/hipaa/for-professionals/privacy/special-topics/de-identification/index.html.

---. "Summary of the HIPAA Security Rule." *HHS.gov*, US Department of Health and Human Services, 26 July 2013, https://www.hhs.gov/hipaa/for-professionals/security/laws-regulations/index.html.

Philosophy of Medicine - Wikipedia. http://en.wikipedia.org/w/index.php?title=Philosophy_of_medicine&oldid=774113132.

Accessed 20 Feb. 2018.

"Storage Gets a Dose of Medical Data - Storage Technology Magazine." *SearchStorage*, http://searchstorage.techtarget.com/magazineContent/Storage-gets-a-dose-of-medical-data. Accessed 20 Feb. 2018.

Xie, Linda. "A Beginner's Guide to Ethereum Tokens – The Coinbase Blog." *The Coinbase Blog*, The Coinbase Blog, 22 May 2017, https://blog.coinbase.com/a-beginners-guide-to-ethereum-tokens-fbd5611fe30b.