A Blockchain-based Framework and Process Guide for Intelligent Exchange and Use of Health Information in Low Resource Environments

NnaEmeka Chukwu

Supervisor: Prof. Lalit Garg

September, 2024

Submitted in partial fulfilment of the requirements for the degree of PhD in Computer Information Systems (CIS).

L-Università ta' Malta
Faculty of Information &
Communication Technology

L-Università ta' Malta

FACULTY/INSTITUTE/CENTRE/SCHOOL Faculty of Information Communications Technology (ICT)

DECLARATION OF AUTHENTICITY FOR DOCTORAL STUDENTS

(a) Authenticity of Thesis/Dissertation

I hereby declare that I am the legitimate author of this Thesis/Dissertation and that it is my original work.

No portion of this work has been submitted in support of an application for another degree or qualification of this or any other university or institution of higher education.

I hold the University of Malta harmless against any third party claims with regard to copyright violation, breach of confidentiality, defamation and any other third party right infringement.

(b) Research Code of Practice and Ethics Review Procedure

I declare that I have abided by the University's Research Ethics Review Procedures. Research Ethics & Data Protection form code 4014_13012020, 4485_03032020, 4948_17042020, 5086_24042020.

□ As a Ph.D. student, as per Regulation 66 of the Doctor of Philosophy Regulations, I accept that my thesis be made publicly available on the University of Malta Institutional Repository.

□ As a Doctor of Sacred Theology student, as per Regulation 17 (3) of the Doctor of Sacred Theology Regulations, I accept that my thesis be made publicly available on the University of Malta Institutional Repository.

□ As a Doctor of Music student, as per Regulation 26 (2) of the Doctor of Music Regulations, I accept that my dissertation be made publicly available on the University of Malta Institutional Repository.

□ As a Professional Doctorate student, as per Regulation 55 of the Professional Doctorate Regulations, I accept that my dissertation be made publicly available on the University of Malta Institutional Repository.

Declaration of Publications

I am writing to declare publications out of this Thesis work *"A Blockchain-based Framework and Process Guide for Intelligent Exchange and Use of Health Information in Low Resource Environments"*. Notably, I adapted parts of sections 2.3 and 3.1.2 of this Thesis to publish the paper *"Systematic Review of Blockchain in Healthcare: Frameworks, Prototypes, and Implementations"*. Similarly, I also improved contents from sections 3.2, 3.3, and 4.1 to publish in *"Standardizing Primary Healthcare Referral Datasets in Nigeria: Surveys, Form-reviews and FHIR profiling"* and *"Digital health solutions and state of interoperability: Sierra Leone's Landscape analysis"*. Also section 2.4 was expanded to publish *"Scaling up a decentralized offline patient ID generation and matching algorithm to accelerate universal health coverage: Insights from a literature review and health facility survey in Nigeria"*. A complete list of publications from this PhD is listed:

Published articles

1. Chukwu E, Garg L, "Systematic review of Blockchain in Healthcare: Frameworks, Prototypes, and Implementations" in IEEE Access, 2020, doi: 10.1109/ACCESS.2020.2969881 [Google Scholar Citations :237]
2. L. Garg, E. Chukwu, N. Nasser, C. Chakraborty and G. Garg, "Anonymity Preserving IoT-Based COVID-19 and Other Infectious Disease Contact Tracing Model," in IEEE Access, vol. 8, pp. 159402-159414, 2020, doi: 10.1109/ACCESS.2020.3020513. [Google Scholar Citation: 233]
3. Chukwu E, Gilroy S, Oyeyipo A, Addaquay K, Jones NN, Karimu VG, Garg L, Dickson K "Formative study of mobile phone usage for family planning amongst young people in Sierra Leone: A global systematic survey" in JMIR Formative Research. doi: 10.2196/23874 [Google Scholar Citation: 8]
4. Chukwu E, Foday E, Konomanyi A, Wright R, Garg L, Smart F "Sierra Leone's health facilities' electricity, computing-hardware, and internet infrastructures: Field mapping" in JMIR Medical Informatics. doi:10.2196/30040 [Google Scholar Citation: 7]
5. Chukwu E, Garg L, Obande-Ogbuinya N, Chattu V "Standardizing Primary Healthcare Referral Datasets in Nigeria: Surveys, Form-reviews and FHIR profiling" in JMIR . 05/03/2021:28510, doi:10.2196/28510 [Google Scholar Citation: 4]

6. Chukwu E, Foday E, Konomanyi A, Wright R, Garg L, Smart F "Digital health solutions and state of interoperability: Sierra Leone's Landscape analysis" in JMIR Formative Research. 2022, doi:10.2196/29930 [Google Scholar Citation: 3]
7. Chukwu E, Ekong I, Garg L, "Scaling up a decentralized offline patient ID generation and matching algorithm to accelerate universal health coverage: Insights from a literature review and health facility survey in Nigeria" in Frontiers in Digital Health. 2022, doi:10.3389/fdgth.2022.985337 [Google Scholar Citation: 3]
8. Chukwu E, Garg L, "RegistryChain: A Framework for tokenized & pandemic-ready shared-data custodianship" in IEEE Transactions on Evolutionary Computation. preprint [UNDER REVIEW]

Published Book chapters

1. Chukwu E., "The role of digital ID in healthcare," in HealthTech law Regulation., pp. 167–192, 2020, doi: 10.4337/9781839104909.00018., Edward Elgar Publishing
2. Chukwu E., Garg L., Zahra R. "Internet of health things: opportunities and challenges," in Artificial Intelligence and the Fourth Industrial Revolution., pp. 27 , 2022, doi: 10.1201/9781003159742., Jenny Stanford Publishing

Patents

- Garg L, Chukwu E, Nasser N, Chakraborty C, Garg G. An integrated secure COVID-19 contact tracing framework using IoT platform with blockchain. Published online 2020.

Acknowledgements

The completion of this PhD thesis marks the end of an incredible journey, one that would not have been possible without God and the support, guidance, and encouragement of many individuals and institutions.

First and foremost, my deepest gratitude go to Prof. Lalit Garg, who started as supervisor, turned mentor, and now friend. He spent countless hours meeting, providing direction, reviewing publication drafts, providing insight to shape the thesis. I also want to specially thank the Maltese Government, for without the fee waiver, and this journey would have been much more difficult. Special thanks to the Faculty of ICT, University of Malta for providing conducive environment for the research.

I am deeply thankful to amazing public health and clinical practitioners who provided much needed guidance and research collaborations. Special thanks to Prof. Edith Ogbuinya, Dr. Iniobong Ekong, Dr. Kim Dickson, and Dr. Patricia Mechael .

Prof. Joshua Ellul shaped my early exposure to the world of "blockchain and smart contract". And working with Matthew Scerri at KPMG also exposed me to real-world blockchain and cryptocurrency. I must also thank members of the Doctoral committee who provided detailed feedback that helped improve the overall Thesis. Finally, I cannot fully express my appreciation to Martha Chukwu for holding the home-front while I was away, taking care of Peniel and Alex by herself. She 'deserves some accolades' for this.

Abstract

Enterprise software systems integration can be simple or complicated depending on the number of components and predictability of component interactions. Designing enterprise software systems with many adaptive components that learn as they interact is not easy. Software systems are often designed as function-specific systems that mimic user concerns modeled around organizational structure and communication patterns. Even a single and simple enterprise now has multiple integrated applications. Traditional integration styles are file sharing, shared databases, remote procedure calls, and messaging. These traditional approaches often require a trusted and centralized access-issuing database owner for multi-stakeholder enterprise systems. In the last decade, a new *trustless* software integration pattern has been facilitated by blockchain. Enterprise blockchain frameworks have been developed, yet practical use cases are few. Use cases still require domain data standardization, token modeling, and interface for regulatory intervention while preserving participants' privacy.

This Thesis investigates enterprise integration using the Health Information Exchange (HIE) use case whose value proposition to healthcare stakeholders is well established. Governments, software vendors, non-profits, and private players traditionally perform the central HIE intermediation role. These intermediation efforts faced many bottlenecks. One main challenge is exchanging large numbers of structured, unstructured, and standardized datasets and terminology sets. This complexity has, over the years, resulted in many healthcare data standards.

Similarly, different practitioner licensing regimes further complicate data controls and custodianship. Governance, ethics, privacy, security, and lack of transparent economic motivations limit these efforts. Consequently, many of the envisaged values of HIE remain unrealized, as further exposed by the recent pandemic responses. In light of the above, the thesis posits the following hypotheses:

1. A decentralized and oracle-friendly software integration pattern will facilitate stakeholder shared *data custodianship* and *economic value*.

2. *Standardized* healthcare registries and repositories on *permissioned shared ledger* will facilitate trustless Health Information Exchange (HIE).

A novel *Regulated-Federated-Decentralized (RFD)* framework is proposed. The RFD framework makes available features that allow data oracles (or regulators) to guarantee trusted data through any combination of real-world approval (or chain of approval) processes on software code. The RFD framework will help cryptographically enforce shared data custodianship and economic value. The structure, privacy, and security of exchanged data, as defined by sharing parties, can be preserved. The RFD framework was shown using the reference implementation, RegistryChain. RegistryChain demonstrates a typical multi-party healthcare registry and repository exchange. The ontology of the proposed framework includes the *Terminology concept* attribute, *Practitioner* list and licensing, and the *health facility* database of services. A novel conceptual Tokens standardized to the Ethereum blockchain format (ERC20) compliant token economic model was designed and presented. The reference algorithms for 1)organization-join token minting, 2)qualifying integrity token awards, 3)health asset transfer (commit and read) tokens management, and 4)tokens transfer were developed. The model facilitates organization stake, structured asset transfer, token transfer, and network operation incentive.

RegistryChain was evaluated by simulating two data types - aggregate (or discrete) data often used for public health purposes and individual-level data often used for individualized patient care. The RegistryChain ontology was checked and validated for consistency. The aggregate data was generated and structured in Fast Healthcare Interoperability Resource (FHIR) format to illustrate and simulate how actual public health data was used. The patient-level data, representing clinical data, was also committed to the test blockchain. The simulation shows that the proposed framework uses less network, computing, and storage resources than equivalent frameworks while allowing for transparent token ownership.

Current systems for HIE rely on either a centralized exchange provider, Patient-controlled, Federated, or the emerging decentralized models. Neither of these architectures has proven optimal for stakeholder-shared custodianship with support for regulatory oversight. Also, current models require sharing institutions (or the intermediary) to be online at runtime for successful sharing.

RegistryChain is the first framework to my knowledge:

- to facilitate multi-stakeholder registry information change management with no intermediary;
- that is privacy-preserving and facilitates transparent shared health data and shared economic value like token amongst stakeholders;
- where communicating parties do not have to all be online at runtime; and
- will facilitate data and service integrity, quality, and ease of audit.

Implementing the output of this work in a health system, has the potential to increase healthcare service quality and data quality through automated accountability in the health system. This will benefit patients, practitioners, health organizations, Electronic Medical Records (EMR) vendors, and development organizations. Opportunities to extend this Thesis in future were equally outlined.

Contents

List of Figures

List of Tables

List of Abbreviations

1 Introduction

Software systems are designed to mirror organization structures, the people, and the workflows there in [5]. Melvin Conway, in his paper published in 1968, "How do committees invent?" concludes with the now-famous *'Conway law'* that *"Organizations which design systems are constrained to produce designs which are copies of the communication structures of these organizations"* [5]. The complex theory paradigm further expands this and considers software enterprises from the *number of components* and *predictability of interaction* of these components [6]. The theory first categorizes a software system as either simple or complex. "System" is used loosely here, as it can be a combination of many different standalone software or software systems. A simple system will have few components, and a complex system will have many components. Additionally, static systems with predictable interactions are considered simpler than an equivalent system with the same components but dynamic and less predictable interactions.

Health systems are an example of a complex system because it is adaptive, with many stakeholders and components that learn and adapt as they interact. Different interoperability components, as in Figure 1.1, further illustrate this complex interaction. Considering healthcare as the use case of interest, one way to view this interaction complexity is by the number of components or the unpredictability of each interaction. For instance, in addition to data sources, stakeholders often have to incorporate domain standards for syntax and terminologies, business workflow, and governance in interoperability models. The complexity and fragmentation in healthcare are better visualized from the lenses of recent digital health investment redirections by Google Health, Health Habit (Apple), Health Vault (Microsoft), Amazon care, and others [7].

1.1 Problem statement

Traditional software integration approaches have their advantages and disadvantages. Notably, traditional software integration styles either need a trusted central intermedi-

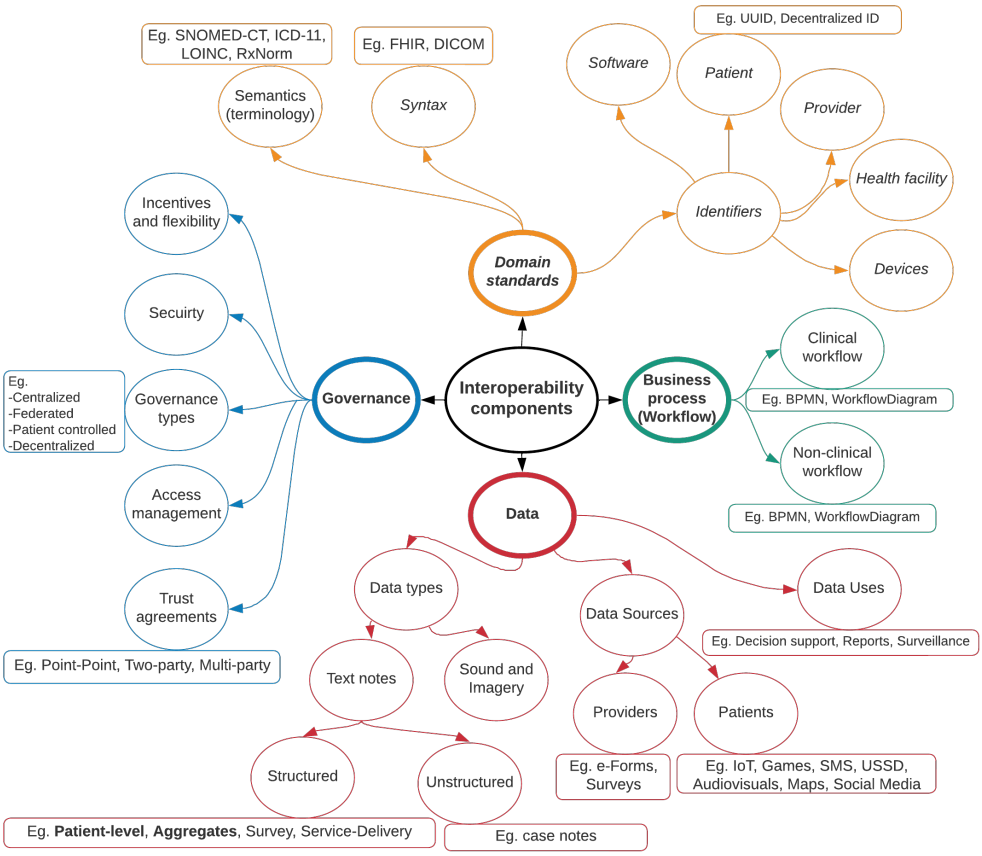


Figure 1.1: Health system interoperability components (self-drawn)

ary or they 'blindly' trust content providers while making Remote Procedure Call (RPC) or sometimes Application Programming Interface (API) calls [8]. The trust in a central authority can be in the form of authentication and service authorization providers. Organizations have traditionally mitigated the need for an intermediary by creating a holding company (or organization) co-owned by transacting parties. A notable example is the Nigeria Inter-Bank Settlement System Plc (NIBSS) "..incorporated in 1993 and owned by all licensed banks including the Central Bank of Nigeria" [9]. Another example is the Indianan Health Information Exchange (IHIE), a non-profit HIE organization with trustees drawn from diverse healthcare stakeholder groups in New York [10].

Individual and organizational territorial tendencies can manifest as 'power-play' conflicts related to who retains the 'information exchange intermediary' role as a 'power control point'. While different software integration patterns for information integration and exchange have been proven, 'information siloes' remain the norm in several multi-stakeholder

sectors, particularly healthcare. When a stakeholder deliberately sabotages rules for information sharing resulting in information-blocking, care coordination is often disrupted [11]. As a result, the many benefits of care coordination are not realized. The drivers of these tensions may vary from financial incentives to privacy concerns and regulations. In the Federated information-sharing model, where data is held and shared on demand is impacted if sharing party is not available at the time of information sharing.

Blockchain, a new form of shared ledger software integration style, facilitates decentralized intermediation. However, limited evidence still exists of its uptake. There is a need to design a comprehensive model that transparently facilitates *shared custodianship* for *services*, *data*, and *economic values* at the point of interoperability.

1.2 Hypotheses and Research questions

In addition to classifying Health Information Exchange (HIE)s by control and governance, any of the HIE models can be administered and managed by either the government, private sector, EMR vendors, Non-profit organizations, or communities. As a result, HIEs can further be viewed as directed, queried, or multiple. Despite these many options, there is a need for an HIE that meets all the criteria of being decentralized in stakeholder service and data control, optimizing regulation, and supporting Patient-centeredness.

Standards, privacy concerns, data ownership, data retention, database control, payment, and incentive arrangements continue to highlight how expensive trust can be. Data-sharing agreements have traditionally been used to solve these problems with marginal success. However, the multidimensional nature of healthcare makes it difficult to cover every case efficiently in any data-sharing agreement [11]. The often undesirably slow agreement changes further compound this problem. In addition, there are 'Oracles' or *trusted-data-sources* in healthcare – i.e., pharmacy, insurance, licensing regulators, and more. And not all can (or are willing to) act as information intermediation providers. Even if they are all willing to act as trusted intermediaries, it may not be practical, as there can only be one intermediary per information-sharing network.

1.2.1 Research Hypotheses

This thesis aims to design and present an optimal framework for exchanging standardized and regulated health information and service while distributing data processing and stor-

age custodianship. Data or data processing custodianship refers to the organization that owns or manages the database providing the data or the service. *Economic value* is the monetary value arising from data sharing or processing.

The two research hypotheses are:

1. A decentralized and oracle-friendly software integration pattern will facilitate stakeholder shared *data custodianship* and *economic value*.
2. *Standardized* healthcare registries and repositories on *permissioned shared-ledger* will facilitate Health Information Exchange (HIE) without an intermediation trusted party.

1.2.2 Research Questions

The following research questions were derived from the hypothesis that a decentralized and oracle-friendly integration pattern will facilitate shared data custodianship, shared services control, and shared economic value:

- What software integration model does not need a central intermediation authority (or availability of data generating institution) for data access?
- How can multiple parties transparently share data economic values without relying on a central intermediary for data sharing?

Similarly, based on the hypothesis that standardized healthcare registries and healthcare repositories will simplify technical HIE, The following research questions were posed:

- What is the current global state of the art of HIE?
- What healthcare data standards are mostly used for semantics and syntax to facilitate HIE?
- What integration architecture allows stakeholders to keep data and data-processing custodianship in a transparently regulated environment?

Clinical concepts are dictionaries of clinical terms curated by experts for cross-communication within a healthcare domain or practice. Similarly, health facilities are hospitals or clinics where care is provided. The practitioner is the healthcare provider.

1.3 Proposed Solution

The proposed Regulated-Federated-Decentralized (RFD) model combines the great features of Federated and Decentralized software integration models. It facilitates oracle-data-management like regulatory approval, license management, and sanctions enforcement. The reference implementation of RFD based on HIE use case codenamed RegistryChain was designed and implemented. A conceptual high-level, organizational, network, and software architecture component was designed and presented. RegistryChain help facilitates transparent, healthcare data economic value sharing. The RegistryChain ontology was designed and tested using the Prodégé ontology modeling program, including these registries and their data structures. Fast Healthcare Interoperability Resource (FHIR) is now popular, and the most used healthcare standard for syntax structuring. Systematized Nomenclature of Medicines Clinical Terms (SNOMED CT) is the healthcare terminology with the most extensive clinical concepts and relationships; others investigated are the Normalized list of US clinical drugs (RxNorm), Logical Observations Identifiers Names and Codes (LOINC), International Classification of Diseases (ICD)11, and Digital Imaging and Communication in Medicine (DICOM) [12, 13, 14, 15, 16]. The RegistryChain blockchain nodes and smart contracts were designed for aggregate healthcare registry services and FHIR structured and shared IPS ontology.

1.3.1 The RFD network architecture

The proposed RFD software integration architecture is modeled after a permissioned blockchain whose novelty is to facilitate the storage of shared health terminologies like services (also known as registries) and shared health data (also known as repositories) without an intermediary. It is based on the zero-knowledge-proof concept where parties share enough to prove presence, absence, or other concepts without sharing more than is necessary for the proof [17].

The architecture uses tokens to facilitate transparent economic value distribution and network incentives. The high-level logical representation of the RFD architecture is shown in Figure 1.2. Based on the RegistryChain HIE reference model, other concern-specific views of the architecture like the UML state diagrams, transaction sequence diagrams, token economics Business Process Modeling Notation (BPMN), and detailed Network diagrams followed. The Web Ontology Language (OWL) for RegistryChain data model is presented as part of the proposed framework. The reference network is composed of distributed nodes providing one or more functions.

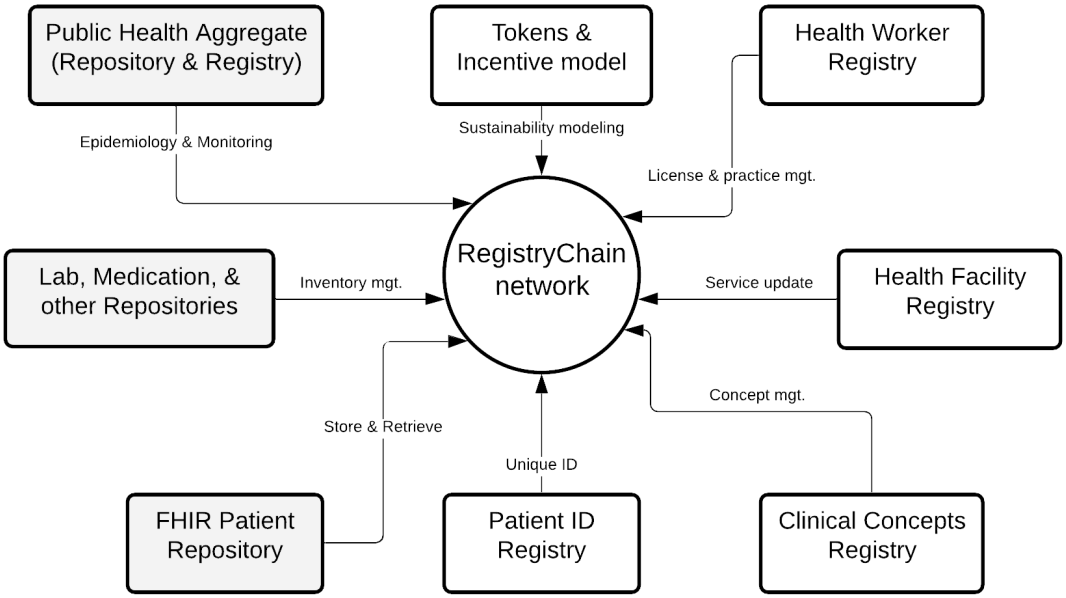


Figure 1.2: RegistryChain logical architecture of RFD (self-drawn)

1.3.2 Reference software architecture

The user concerns of priority were grouped from the architecture into Participants, Functions (the basis for Transactions), Assets, and state attributes of the assets. In Reference RFD architecture for health registries change management (RegistryChain), Participants are Organizations or Patients, Practitioners, and NetworkAdmins. The Assets are also tied to a particular participant with an established role in the network. Each participant's role is mapped to their function (or transaction). Assets have been chosen as a specific use case to test the framework. Asset attributes are standardized first to FHIR and extended as necessary. The Asset attributes constitute state property values, which change following transactions committed to the blockchain world state. In the RegistryChain HIE architecture, the multi-CA architecture model for identity generation, updates, and revocation was used. The smart contract policy for creation and updates is well-defined. The RegistryChain's tokens can be generated, staked, transferred, or used for service. Data storage, processing, query, size, and node uptime determine token earnings or usage on a RegistryChain network. RegistryChain also provides data aggregation support needed in healthcare for monitoring and epidemiology support. The software architecture is designed to conform to the Richardson maturity model level three for Representational State Transfer protocol (REST) API [18]. A subset of foundational RegistryChain services

is listed:

- FHIR validator and conformance service,
- Clinical concepts update manager,
- Health facility attribute and service listing manager,
- Health workforce attribute and license manager,
- Organization network membership manager,
- Patient Identity manager,
- Transaction endorsers,
- Transaction orderer,
- Storage providers,
- Cross organization nodes (anchors),
- Certificate Authority provider and manager,
- Client query relay-er

1.3.3 Standards & Ontology

RegistryChain model used the FHIR Revision 4.0.1 standards core profiles naming nomenclature for modeling the datasets that make up the RegistryChain ontology. The reference OWL ontology includes PractitionersRegistry, OrganizationRegistry, TerminologyRegistry, and International Patient Summary (IPS), all represented in FHIR bundles (a kind of array of JSON) format.

The RegistryChain ontology in the architecture has a novel on-chain terminology and repository services. To my knowledge, no such architecture currently exists. The process for updating and use of these terminologies in practice currently varies greatly. The RegistryChain model allows for flexibility in who can post an update to the terminology and supports smart contract approvals. This flexibility guarantees continued regulation and quality checks. An attribute's cardinality or concept relationship in an ontology may be updated or approved for replication through the traditional approval (or chain of approval) process(es). This workflow is implementable in a smart contract with adequate

permissions. This same workflow can be used for health facilities or other registries.

The reference implementation RegistryChain includes support for a poly-hierarchical registry and meaning-based clinical concepts terminology model like SNOMED CT or RxNorm. RegistryChain replicated this model in the clinical concept on-chain change management model. The model also supports uni-hierarchical and statistical classifiers like ICD. The ICD10 or ICD11 provides a statistical classification of concepts without relationships and is used globally for disease surveillance and reporting.

1.4 Thesis Outline

See Figure 1.3 for sections and subsections of the Thesis that address the different research questions. The remaining sections of this Thesis are arranged next to present the reviews of related work. The reviews systematically searched and discussed the state of evidence of 1) the global state of the art of HIE 2) the barriers to data sharing; 3) the uses, effectiveness, and health outcome of current HIE systems; 4) the governance models used for HIE, the reference ontology; and 5) and how tokens have traditionally been used in healthcare literature.

The next chapter discusses the background and state of the art in software integration and HIE. The three-part study methodology, which includes the approach for the Country-level survey that mapped the state of digital health interoperability in a typical Low and Middle Income Countries (LMIC) is presented next. The network and software architecture design strategy and implementation approach. The Proposed framework chapter details the proposed Regulated-Federated-Decentralized (RFD) framework, which is an oracle-friendly information system integration pattern based on the zero-knowledge-proof feature of blockchain. In addition, RegistryChain was discussed, a reference healthcare ontology use case that can be used to exchange healthcare registries for Practitioners, Health facilities, and terminologies while transparently documenting shared economic value using tokens designed on smart contracts. It also included the network, software, and business architectures. PractitionerRegistry, OrganizationRegistry, and TerminologyRegistry components enabled by RegistryChain are regulatory functions in most health systems. The ontology was checked and validated for consistency. Also, the FHIR resource conformance metric and the Hyperledger fabric Performance metrics were used as primary evaluation schemes. The thesis concluded with a key industry significance, outstanding future works, and wider implications of this framework.

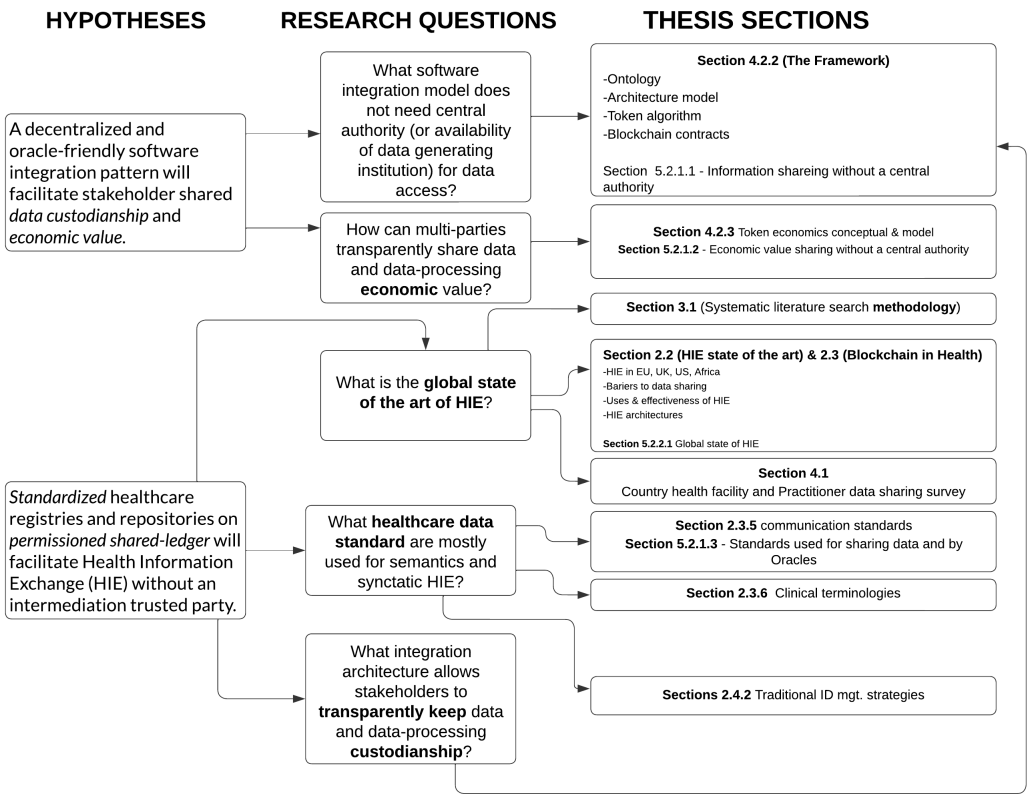


Figure 1.3: Sections addressing different research questions(self-drawn)

2 Background and the State of the art

This chapter first discusses the overall Thesis background, including software and healthcare integration pattern and architectures. The systematic literature review of Health Information Exchange (HIE) was presented next. This was followed by findings on the global state of HIE, which presented barriers to information sharing, uses, and outcomes. In addition, the state of the art of HIE by geographic groupings are also discussed. The emerging decentralized blockchain-enabled architecture and supporting components for HIE application are also discussed in this chapter.

2.1 Background

This section discusses enterprise software architecture and how it relates to user concerns, viewpoints, and views. The enterprise software integration patterns were next discussed. Afterward, the blockchain-facilitated integration pattern was discussed, and then the current healthcare integration patterns and healthcare interoperability components.

2.1.1 Enterprise software architecture

Enterprise architecture involves analyzing, planning, designing, and eventually implementing an enterprise. The architecture usually includes a collection of diagrams and relevant supporting documents describing a software system [19] [20]. Each diagram presents a view of the whole system ignoring everything else. Enterprise-grade software systems are complex and challenging to design and deploy because they combine the challenge of understanding a business domain (like healthcare) with managing a team of software and sometimes hardware teams through several phases of development. Time-to-market pressure further compounds this challenge resulting in failures, canceled projects, and cost overruns. Many completed projects sometimes contain fewer capabilities than promised and are hence considered unsuccessful [20]. Modifications of finished software

products are expensive and can create more defects. Enterprise-grade systems will have persistent relational or file-based databases or both. They will be connected to different user interfaces. The user interface, for instance, can be a command line, or Graphical User Interface (GUI), or the data types can be binary, text, imagery, or multimedia.

Remote access is supported, and third-party packages are used. Distributed computing and storage have dramatically increased software systems' complexity, resulting in rising user expectations and requirements. User expectations are also continually changing, resulting in the constant evolution of software products. Different software systems can be categorized thus: legacy system software; in-house software; commercial off-the-shelf software; freeware (or public domain software); shareware; prototype software. Software systems developed without adequate architectural artifacts are considered ad hoc. Ad hoc software systems result in poor (or no) documentation; they are challenging to maintain and a nightmare to upgrade. Staff turnover can expose the inadequacies of ad hoc software systems, as the organizations often need help with an unmodifiable software system(s) [21]. Architecture-driven software development approach can address many of these challenges. Gamma et al. first introduced the concept of software design patterns in their timeless book "Design Patterns: Elements of Reusable Object-Oriented Software" [22]. Also, Martin Fowler extended this idea to detail design patterns in software development [19].

2.1.2 Concerns, Architectures, Viewpoints, and Views

Different aspects of a software system are of interest to different categories of stakeholders. For instance, some stakeholders may be interested in the high-level enterprise view. Others will be interested in the unit, departmental, or organizational views. Others may be interested in the domain or detailed software views. The IEEE' recommended practice for architectural description of software-intensive systems' defined a view as "a representation of a whole system from the perspective of a related set of concerns" [23].

On the other hand, concerns are those interests that relate to the system's development or critical aspects to one or more stakeholders. Typical concerns are security, maintainability, scalability, reliability, or performance. Similarly, "A view represents a partial aspect of a software architecture that shows specific properties of a software system" [24]. Similarly, "viewpoint is a specification of the conventions for constructing (the 'how'), interpreting and using a view. Viewpoints can be seen as templates for developing each view by establishing its purposes and audience and the technique for its creation and analysis"

[23]. The conceptual framework adapted from ISO42010 is shown in Figure 2.1.

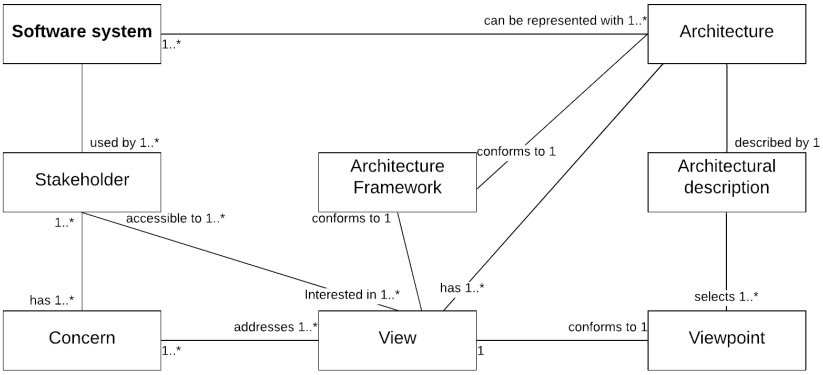


Figure 2.1: Representation of concerns, views, and viewpoint relationships

Similarly, an architectural model often has multiple views. Each architecture view addresses one or more stakeholder' concerns'; thus reducing complexity by allowing each stakeholder to focus on their concern(s). The different Universal Modeling Language (UML) [25] notation diagrams, the Business Process Modeling Notation (BPMN) [26], and the System Modeling Language (SysML) [27] notation can all be seen as viewpoint frameworks. On the other hand, the specific healthcare domain UML, BPMN, or SysML diagrams are the views that constitute the artifacts of interest in the software architecture.

2.1.3 Enterprise software integration patterns

Most healthcare applications of value often get integrated with other applications. A typical health system or enterprise often comprises many custom-built, third-party, legacy, multi-tier systems running on different platforms and operating systems. They sometimes use different data formats. A reason for fragmentation is that business applications are difficult to engineer, and a one-stop solution for every business scenario is almost impossible. The inability to use one all-in-one application is a well-known conundrum, even within a sector like healthcare. Organizations are constrained to design systems that mimic their current workflow, communication pattern, and organization structure [5]. Hence, function-specific applications often provide the flexibility to select the 'best' immunization system, the 'cost-effective' insurance claims management system, and much more. As such, there is an increasing preference for function-specific applications by vendors and product owners. However, delineating these functions is difficult. For instance, a

claims issue may be considered a point-of-care problem or a billing problem. These many system design approaches blur and ambiguates the business boundaries that the Patient or Practitioners are often not concerned about. Patients and Practitioners, for instance, focus on functions tied to their roles or views of interest. These functions sometimes change depending on the healthcare establishment or the jurisdictions.

Hohpe and Woolf classed integration patterns as file transfer, shared database, remote procedure calls, and messaging, with six project scenarios types: Information portal, Data replication, Distributed business process, Business to business integration, Shared business function, and Service Oriented Architecture (SOA) as in Figure 2.2 [8]. In the *information portal* scenario, information from different sources can be queried from one source (or API aggregator). Two or more independent software systems can also use the *data replication* scenario for information sharing. Two or more business applications can share a service (or a well-defined business function), e.g., authentication using *shared business function*. Business applications can use service discovery, negotiation, and service directory listings for sharing and managing common services using *SOA*. In less common cases, businesses can directly collaborate in a two-way *business-business* integration pattern. For redundancy, business functions and resources may be distributed using the *Distributed Business Process* integration pattern.

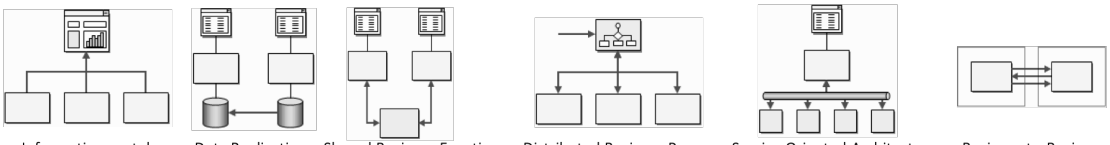


Figure 2.2: Integration project scenarios according to Hohpe and Woolf [8]

Traditionally, software integration can be facilitated by file sharing, shared database, remote procedure calls, or messaging [8].

2.1.3.1 File transfer integration style

File storage remains an integral part of any enterprise operations and operating system. Important decisions in using the file integration approach are the data format (e.g., text-based files or XML), when (time) to generate/consume [8]. File transfer is best suited for integration that happens less frequently, like nightly or monthly. A key advantage of the file transfer approach is that application components need no knowledge of the internals of the communicating applications. They individually agree and negotiate formats, locking (or agree on timing) without extra integration tools and packages [28]. File transfer can

use any combination of architecture, client-server, server-server, or client-to-client [28]. However, the developer will have to do extra work to achieve integration for file transfer. The biggest challenge with file-based integration is managing inconsistencies arising from multiple updates between set synchronization times [8]. One way to minimize inconsistencies is to reduce synchronization time or create files as changes happen. The drawback with producing and processing large amounts of files quickly is that it easily gets resource intensive.

2.1.3.2 Shared database integration style

Speed of data processing amongst multiple independent platforms in an enterprise is often the primary goal. In many enterprise applications, data staleness is not tolerated; where this happens, it leads to errors and systemic user distrust in data. Rapid and frequent updates reduce inconsistency to the barest minimum, ensuring they are properly handled. However, rapid changes come with their problem, especially if updates happen in rapid succession inconsistently [8]. Databases help deal better with "semantic dissonance" related to different ways of looking at the same data. Shared databases leverage SQL-based relational databases for transaction management [29]. The file-format problem and 'semantic dissonance' become less of an issue in this approach. The drawback with the shared database is the suitable initial design, like modeling the unified schema that will address the needs of the different enterprise applications. On the other hand, the political challenge of choosing a shared database's custodianship can sometimes make it unsuitable for multi-organization (or inter-department) operations. These "human-political" conflicts and competition between departments or organizations often exacerbate issues [8]. As multiple applications use a shared database, a major drawback an enterprise can experience is performance bottlenecks resulting from each application locking others out of data at runtime. Though recent advances in DataBase Management Systems (DBMS) ensure data integrity, efficient and highly scalable shared databases using record-locking (or even field locking). Such a system comes at a financial cost beyond the reach of average users. One solution will be a distribution of databases with offline capabilities. This then brings the synchronization issues and conformance to Atomicity, Consistency, Isolation, and Durability (ACID) properties to the table. Shared databases are also limited because they provide a large unencapsulated data structure (limited separation of concerns or little modularity), making responsiveness difficult [8].

2.1.3.3 Remote Procedure Call (RPC) integration style

Enterprises sometimes need to responsively share processes (in addition to data) between departments and across institutions. Changes in data often trigger activities and processes with them. An application must know enough of the internal processes of a communicating application. In software architecture, this problem is resolved through encapsulation (hiding details). Remote Procedure Invocation principles ensure that every application manages and maintains the integrity of its own data through the principles of encapsulation (or separation of concerns) [8]. If an application needs data stored in another application, they ask through a direct read. If an application wants to change the data in another application, they call a pre-defined method (or procedure). A key advantage of this approach is that each application can update and change its internal data and structure without affecting others. Simple Object Access Protocol (SOAP) is a popular Remote Procedure Call (RPC) approach using Xtensible Markup Language (XML) as data format over Hypertext Transfer protocol (HTTP) web service. Representational State Transfer protocol (REST) is a typical RPC architectural pattern based on RESTful principles [30]. These are typical for shared healthcare registries. The method wrappers make dealing with 'semantic dissonance' easier. One drawback of RPC is that some developers sometimes do not consider the differences in network and computing infrastructure performance and reliability. Despite eliminating large shared data structures (as is in shared databases) in RPC model, applications are still tightly coupled. Tight coupling here concerns application development and execution.

2.1.3.4 Messaging integration style

A messaging system should have components that have high cohesion (local work-intensive) and low adhesion (remote light work) [8]. Transformation can happen during messaging between sender and receiver without both being aware of such transformations. One advantage of decoupling is that senders can broadcast messages to multiple receivers [31]. Alternatively, integration can be modeled to send messages directly to one or a few of many possible recipients. Other topologies between this two are also possible. The ability for possible transformations means that communicating applications can have different conceptual internal models [8]. While this introduces the challenge of semantic dissonance, a uniform messaging viewpoint, as used in a shared database pattern, can mitigate this. The added advantage is that applications share data while collaborating behaviorally. The messaging approach ensures that callers can continue while their message is being processed. Though, with messaging, systems do not quite update simultaneously.

Messaging has facilitated distributed systems development, yet access issuance and permission management remain centrally controlled like other integration styles.

2.1.4 Blockchain facilitated software integration style

The software integration patterns described in the preceding subsections assume that a trusted entity manages permission to access resources. However, blockchain facilitates trustless intermediation without a central authority. Blockchain (one form of Distributed Ledger Technology (DLT)) gained prominence after the 2008 global financial crisis. The first mention of blockchain was in the bitcoin white paper by Satoshi Nakamoto [32]. The bitcoin public blockchain was subsequently implemented with bitcoin as its cryptocurrency. The bitcoin blockchain is a public blockchain network because anyone with the desired computing power and network bandwidth can join the bitcoin network. Bitcoin provided a world state of the shared ledger where nodes store synced world state ledger. The ledger is immutable, and data added to the ledger is via a consensus determined by nodes with 51% computing power. Nodes that participate in transactions and append data to the network are rewarded through an incentive in bitcoin cryptocurrency. Transactions can be asset logging on the blockchain or transfer of cryptocurrency ownership. Bitcoin limited the transaction script size and amount of data that can be added to the blockchain ledger and used Proof-of-Work (PoW) consensus algorithm to enforce this while securing the network. Many consensus algorithms have been proposed to address the wasteful energy consumption of PoW consensus, already discussed in the survey paper [1].

The second generation of blockchain networks introduced smart contracts, first used on the Ethereum blockchain network, inspired by bitcoin [33]. Ethereum also uses PoW consensus mechanism, though, without the transaction size limitation, it enforces security by requiring payment for adding data to the ledger in Ether cryptocurrency. A new wave of private (or permissioned) blockchain networks followed, which attempted to solve the wasteful PoW problem that limits blockchain scalability by identifying participants. Hyperledger Fabric, an open-source blockchain by a consortium led by the Linux Foundation, is a leading enterprise-grade blockchain solution [34]. The advent of permissioned blockchain ushered in a wave of blockchain applications in many sectors, including healthcare. Most of these applications are driven by the stakeholder disintermediation, data integration, and scalability of permissioned blockchain solutions.

2.1.5 Healthcare concerns & integration

Health systems are prioritizing quality health service through equitable, effective, efficient, safe, timely, and patient-centered services [35]. Healthcare quality issues are estimated to cost about five million deaths in 137 countries in one year [36]. The global scale of poor patient service delivery is not yet fully investigated. However, in the US alone, healthcare wastes emanating from unnecessary services, inefficiencies, fraud, overpriced service, high administration costs, and missed prevention accounted for $765 billion in one year [6]. At the same time, Patient and population health information is collected as the patients traverse the health system. An average Patient would typically see many different healthcare service providers during a care episode and in their lifetime. Investment in Health Information technology (IT) in the US alone was estimated at $760.2 billion annually in 2024 and expected to have grown to 1.8 trillion by 2030 [37]. Despite these investments and progress in adoption of digital systems, providers and health systems leaders believe interoperability gaps remain. Some due to non-existence of or limited maturity of foundational shared information structures like Master Patient Index (MPI) and Standards. For paper-based systems, getting a complete picture of a patient's health would mean going over hundreds of pieces of paper distributed in a dozen or more locations. Information Communications Technology (ICT) is driving acquisition and information use across health systems. The World Health Organization (WHO), in recognition of the unique place of digital technologies in health systems, championed a World Health Assembly (WHA) resolution for the digitalization of health systems [38].

Despite digitization (and digitalization) progress, the demand for quality data in support of better healthcare decisions both at individual and public health levels is rising. Healthcare data quality concerns border on accurate, concrete (or indisputable), representative, of high integrity, precise, consistent, relevant, complete, and timely data [39]. A complete view of a patient's health information at any time can help address many of these data and service quality concerns. However, healthcare data sources vary with many different stakeholders, interest groups, and related power dynamics. The primary and widely known source of health data is the healthcare practitioner, captured in paper or electronic forms based on interviews, procedures, or inferences. Other sources are also emerging, like patients and electronic wearables. In order to maximize the different data sources and types, coordination and integration are required. Also, healthcare practitioners vary significantly by their specialties and sub-specialties.

An Electronic Medical Records (EMR) at a health institution holds a Patient's medical

records within that institution. The Patient can access their health record in the form of a Patient Health Records (PHR). When a unified store and view of a Patient's health information is required, an Electronic Health Records (EHR) provides the complete view. Evidence abounds that digital-enabled healthcare coordination improves effectiveness, intermediate Patient outcomes, data use, and efficiency, and helps with care quality through improved data quality [40]. Globally, the Corona virus 2019 (COVID 19) pandemic further exposed the importance of care service delivery coordination. Also, the WHA in 2013 passed a resolution for health data standards and interoperability prioritization [41]. Similarly, the US Office for National Coordinator (ONC) of Health IT, in its 2019 roadmap, "Shared Nationwide Interoperability Roadmap", prioritized key components to facilitate interoperability of Health IT system [42].

Despite the many visible values of (and widespread interest in) Health Information Exchange (HIE) systems, successful EHR with multi-institutional patient records is not widespread. The reasons are not far-fetched as Benson and Grieve noted that many health Information Technology (IT) interoperability projects fail (or fail to achieve their original aim) [11]. Healthcare interoperability is **hard** because there are many dimensions and associated components, as already shown in Figure 1.1. An interoperable health system will have to navigate the many different domain standards, Interoperability governance, data frameworks, incentive models, business workflows, and Regulatory constraints.

For instance, a recent study in a metropolitan health institution in Australia identified 69 different clinical registries (or *CodeableConcetpts*) in use [43]. One such registry is the Systematized Nomenclature of Medicines Clinical Terms (SNOMED CT), a widely used polyhierarchical healthcare terminology ontology, which contains over 60,000 clinical concepts with attributes, relationships, identifiers, and descriptions. As healthcare is multi-stakeholder, multi-vendor, and multi-institution, managing cross-institution change management can easily become a nightmare. Updates must be shared amongst institutions that use them with attendant software and governance-related coordination overheads. This complexity exponentially grows considering Disease domains, Patients, Practitioners, Institutions, Devices (Health nodes), data syntax, data formats, and governance models. Currently, terminology service providers develop pieces of software that their users are often required to separately install to receive and manage updates [12, 13, 14]. What this means is that the Metropolitan hospital in Australia described above will have 69 different interfaces, data formats, and software for registry updates management [43]. This problem of registry and repository update management is categorized into two care coordination and collaboration problems:

1. Data sharing
2. Service sharing

The value of health information coordination and use in healthcare delivery has long been established [6, 38]. Benson and Grieve identified how clinical interoperability in EHR is used to facilitate clinical messaging, clinical decision support, orders management, reports, records transfers, retrieval and update of shared records on medical devices, capturing patient data [6]. They went on to highlight decades of failed interoperability projects in the UK, US, and EU. Healthcare interoperability is expectedly complex because of the many specialties and sub-specialties working together to provide care to patients in the Patient's care path. Sharing service control, data custodianship, and economic value remain at the center of many frictions. The main concern of healthcare stakeholders with regards to HIE is to share standardized data across institutions. Some specific concerns are listed here.

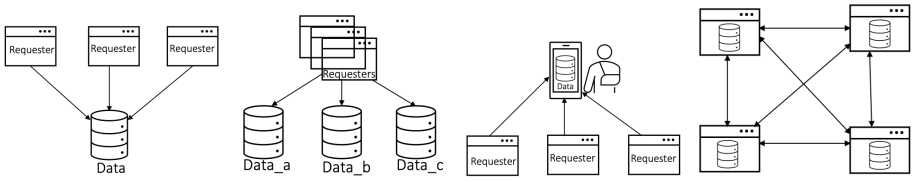
- The **Regulators** want to continue to provide regulatory intervention like registry updates and sanctions.
- The **Patients** wants quality healthcare, better health outcome, and new healthcare services.
- EMR, PHR, EHR **vendors** or stakeholders want to keep *control* of historic *services* and *data* they generate while collaborating.
- The **Health system** needs to generate *aggregates* across domains for planning, disease surveillance.
- **Health system** sustainability needs a data value-sharing model that benefits data generators, owners, custodians, users, and policy decision-makers.
- **Private sector** and **software vendor** want increased revenue profile as a result of scaled care coordination and new healthcare services

2.1.6 Health Information Exchange (HIE)

Health Information Exchange (HIE) is *"..the electronic transfer of clinical and/or administrative information across diverse and often competing healthcare organizations"* [44]. HIE can be used as a verb or noun depending on the sentence context. As a verb, it connotes the action of moving health information between and amongst stakeholders in a healthcare

sector [44]. On the other hand, when used as a noun, it describes a legal entity that facilitates the HIE. Throughout this thesis, HIE will be used in these two contexts, with each context clearly described as such.

Different models have emerged for classifying HIE systems. When classed by the governance level of control, the models can be Centralized, Federated, Patient Controlled, and Decentralized [1, 44]. Each model comes with its pros, cons, preferences, and evidence base. In a Centralized model, one organization is responsible for custodianship and managing the data and access. The Federated model is designed to enable each institution to keep their collected data and make available API endpoints for data query and possibly updates. As the name implies, the Patient-controlled HIE model gives the Patient the power to manage their data through technology service providers directly. In the last decade, there has been active research leveraging Distributed Ledger Technology (DLT) to facilitate health information sharing without a trusted intermediator, see our published review [1]. The logical representation of these four HIE models using three data requesters and data host models is illustrated in Figure 2.3.



(a) Centralised (b) Federated (c) PatientControlled (d) Decentralised

Figure 2.3: Logical view of HIE models (self-drawn)

2.1.6.1 Centralized architecture

The centralized model of HIE governance is one of the most popular as it relies on one stakeholder to host and share the data while managing access to the data. The stakeholder can be either a separate intermediary or a participant in the information-sharing arrangement. The main drawback of a centralized architecture is that the central service provider with the appropriate technical capacity is often a digital health competing vendor. As the amount of data hosted by the intermediary provider increases, so does their power and their competitive edge over others. This scenario leads to the "information blocking" challenge, which the US Office of National Coordinator (ONC) of Health IT flagged in its 2015 report to the US congress [45]. The report explained that "information blocking occurs when persons or entities knowingly and unreasonably interfere with the exchange

or use of electronic health information". Nevertheless, the centralized model is the easiest to audit or regulate. France uses a model of centralized HIE, but adoption has been poor, and the central repository has now been revamped with the intention to drive up adoption [46]. Similarly, some US state government and community-managed HIE models use centralized architecture [10, 47]. The Open Health Information Exchange (OpenHIE) architecture community, popular in developing countries with communities in Ethiopia, Kenya, Haiti, Zambia, and Tanzania, also promotes centralized architectures [48].

2.1.6.2 Federated architecture

The Federated governance architecture allows stakeholder organizations to keep their generated data while data requesters query their managed databases to access resources. Each stakeholder organization determines access controls, whether or not to grant any request to shared data [49]. Regular audits can be used to check agreement compliance, though audits rarely happen in practice. This model also relies on the runtime 'availability' of each organization where the Patient has their data for a complete view of the Patient's health data. Public health and surveillance will require significant work to guarantee that each stakeholder organization remains compliant with syntactic interoperability. Also, each stakeholder organization will host its terminology concepts management system for disease, medications, or identifications. Regulating a federated system is often difficult for many of these reasons.

2.1.6.3 Patient-controlled architecture

When the Patient is in charge of their health records, they are responsible for making the health record available to the provider at the point of care or before the data is needed. Depending on the design, the Patient may have their data hosted by a technology service provider or directly manage their data and share as they see fit [50]. Questions often arise regarding non-literate Patients' ability to manage their health data. The question of how to share an adequate amount of data with the right physician always often arises [51]. The question of inter-dependent conditions, comorbidity, adverse drug reactions, and related relationships complicate patients' general ability to manage their own data. On the positive side, the Office of National Coordinator (ONC) Health Information Technology (IT) is prioritizing Patient-centric healthcare as a strategy for improving Patient care [52]. There is now an increased drive to incentivize Patients or their caregivers when their data gets accessed or used for research [52].

2.1.6.4 Decentralized (or Blockchain) architecture

The last decade ushered in a new wave of opportunity with *zero Trust* frameworks facilitated by blockchain. A decentralized architecture of HIE allows participants to share information through a distributed network and storage schemes. A systematic review of different applications of blockchain for HIE was conducted and published [1]. In the review, global proposals, prototypes, and pilot implementations of blockchain in healthcare were aggregated. Different consensus approaches have been proposed to agree on how to add data to the shared ledger. The leading applications of blockchain for HIE use PoW consensus, which is expensive and has proven non-sustainable. Also, healthcare is a heavily regulated sector: while the decentralized model is promising, many models make regulation difficult. Also, healthcare requires large data terminologies and standards management. Incentive computations are routine, coupled with varying workflows and software preferences. Health stakeholders want enhanced data privacy while fulfilling key regulatory constraints. A key concern is the Regulator's ability to enforce business rules and sanction erring organizations. Many healthcare stakeholders want visibility into how a potential software architecture addresses these concerns, which remain unresolved to my knowledge.

2.1.7 Health Interoperability components

Interoperability is the *"ability of information systems, devices, applications to use, exchange, integrate, cooperatively use data within and across organizational, regional, or national boundaries"* [53]. In healthcare, achieving semantic interoperability requires that both data and the concepts they describe be exchanged in standardized formats. Oftentimes, custom and localized contents in the form of ValueSets are necessary for healthcare service delivery. Uniquely identifying Patients, Practitioners, Health Institutions, Devices, and Applications is equally crucial for care coordination and interoperability. Also, a shared repository accessible to parties in the care continuum, the privacy, and security of shared content remain essential components of healthcare interoperability. Healthcare incentives and service tokens for shared value and trust are also vital for interoperability value. The last and sometimes considered most critical interoperability component is the ability to use aggregate public health and epidemiology data.

2.1.7.1 Standards: Data structure and FHIR

The healthcare domain standard that facilitates healthcare interoperability operates at the application layer of the OSI reference model and helps programmers take maximum

advantage of a distributed healthcare system. Standards can be seen as the product of an organizational process that promotes reusability and interoperability. It can take many forms starting from an ad hoc integration and documentation between two sub-systems or across projects, organizations, industries, or geography. Based on how they come to be, there are three ways standards may originate [54].

- Formal standard
- De jure standard
- De facto standard

Formal standards are those that one or more accredited standards bodies have adopted. De jure standards, on the other hand, are those mandated by legal authorities. At the same time, de facto standards are those that result from widespread usage. De facto standards often result from vendor market dominance and, in other cases, due to popular usage.

Given the critical role semantics and syntax play in healthcare, one would expect unified semantics and syntax for the information in healthcare. Instead, there are many mature sets of often competing Electronic Health Records (EHR) standards syntax and terminologies [55]. Similarly, information-sharing models require persistent storage for storing shared health records. The storage schemes, access control, privacy management, and security determine the success of integration endeavors. Health Informatics authors often class health communication standards into two main categories [11, 45]: 1) the dialect of the content (semantics), 2) and the structure of the message (syntax).

Despite these two broad categorizations, Healthcare standards can also be classed as Vocabulary (Terminology), Content, Transport, Privacy, Security, Identifier, Reporting, Process flow, and architecture. The leading standards for content structuring and communication (syntax) are Health Level Seven (HL7) v2, v3, Clinical Document Architecture (CDA) [56], and FHIR [57]. Transport standards address the push or pull strategies format for HIE. Medical equipment manufacturers widely use the Digital Imaging and Communication in Medicine (DICOM) standard for medical imaging, archiving, and related information [16]. The HL7 FHIR is popular because it uses web standards for data interchange, just like REST, identifying everything as a resource, structured in JavaScript Object Notation (JSON) or XML or other related web formats [57].

The ISO/TS 18308 specification details requirements for the content structure, how a health record is organized, encapsulation, or portability. It further describes secondary users, archiving structure, and data organization, including structured and non-structured data and Clinical and administrative data. Standards also cover types and forms of data, including what type is supported and which data forms reference data. For modeling health software systems, ISO/EN13606 [58], HL7 Reference Information Model (RIM) [56] and OpenEHR [59] are the leading standards.

2.1.7.2 Standards: Terminologies and ValueSets

The leading semantic standards are the Systematized Nomenclature of Medicines Clinical Terms (SNOMED CT) that provides a comprehensive ontology of clinical concepts[11], the WHO for terminologies classification International Classification of Diseases (ICD) [60], Logical Observations Identifiers Names and Codes (LOINC) for identifying health measurements and observations [13], Normalized list of US clinical drugs (RxNorm) used to normalize clinical drugs across vocabularies [15]. In addition, there are other national-level terminologies like medicines formulary.

A Registry (or CodeSystem) is any framework that publishes a list of codes or the rules for generating codes and for what purpose. Many code systems are used in healthcare to structure terms, ontologies, and data. These codes can range from simple gender groups to a complex diagnosis set of concepts. Clinical registries are used to structure clinical concepts better and keep a consistent description, naming, relationships, and codes. In healthcare, codes can be generated for the diagnosis, procedures, laboratory tests, billings, and identifiers. Figure 2.4 from my book chapter gives examples of some coding frameworks used to manage terminologies of popular domains in healthcare [61]. The role of SNOMED CT, LOINC, RXNorm, ICD, DICOM terminologies has already been discussed.

In HL7 FHIR, *CodeSystems* FHIR resources are used as master catalogs of individual code systems in healthcare. Similarly, simpler non-formal code systems use *ValueSet* FHIR resource. The *CodeSystems* resource supports different types of terminology systems thus:

- **Classification terminology** - like self-created gender classification table, a simple table of codes and their description or *ValueSet*.
- **Uni-hierarchical terminology** - like ICD-11 [14], where codes can have either a parent or a child.

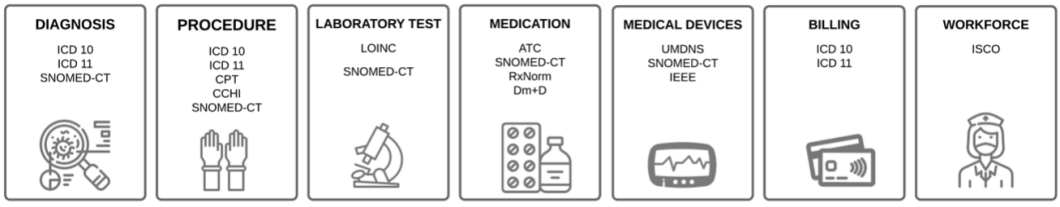


Figure 2.4: Code system groupings with example terminology frameworks [61]

- **Poly-hierarchical terminology** - like SNOMED CT, [12] where codes have multiple parents and multiple children, all interrelated.

2.1.7.3 Patient Identifier

Uniquely identifying Patients, Practitioners, Health Institutions, and related objects remains crucial to effective information integration. See Figure 2.5 for entities whose unique identification is essential in health system designs from my recent book chapter survey [61]. Healthcare increasingly requires more and more integration because of the many different stakeholders and services involved in patient care.

Figure 2.5: Entities for identification within a healthcare system [61]

Patient identifiers are crucial to creating and updating existing records across institutions. The uniqueness of a patient Identifier (ID) ensures the validity of operations on existing records. Within a health institution, the use of a health facility-specific Medical Record Number (MRN) can suffice. However, identity management becomes complicated across multiple institutions. Timothy et al. broadly categorized patient ID into two [62]: 1) a unique code or set of codes designed to identify a patient uniquely. 2) an aggregate of demographic and related attributes used to describe a patient uniquely, e.g., sex, name,

or date of birth. When care happens across institutions or even departments, the need to match a patient to their shared records becomes even more complicated. In practice, health institutions use different schemes to identify patients uniquely. Some use national ID, others hospital MRN, Master Patient Index (MPI), or other functional ID schemes as in my already published work [61]. These options and relevant use cases were extensively discussed in the published book chapter *The role of digital ID in healthcare* [61].

2.1.7.4 Practitioner Identifier

Healthcare practitioners deliver preventive, curative, promotional, or rehabilitative healthcare services to individuals, families, and communities. Many healthcare service providers exist (e.g., physicians, dentists, pharmacists, Laboratory Scientists, Nurses, Midwives, Community Health workers, volunteers, receptionists, informaticians, and social workers). In the FHIR parlance, they are Practitioners, which will be how they are categorized in this thesis. Each category of health workforce has a relevant regulating agency with membership, certification, and licensing necessary for continued practice. Practice licensing information is often needed for cross-institutional care coordination and across jurisdictional and geographic boundaries. The WHO also promotes this - "the establishment of a national health workforce registry is essential for strengthening national health systems at all levels" [63].

2.1.7.5 Health Institution Registry

There are currently no universally agreed-to standards or datasets for health facility data collection and update. The widely used HL7 FHIR interchange standard uses the *Organization* resource as the closest resource for this purpose. Nigeria, for instance, established a national *Health Facility Registry (HFR)* for one-stop health facility data management [64]. Likewise, other countries, particularly in sub-Saharan Africa, have similar initiatives. The *HFR* will add significant value to care coordination by making available a transparent mechanism for regulatory organizations to review and approve suggested changes to the *HFR*. For instance, a near-real-time availability of service listing to the public can greatly improve efficient service utilization. However, limited evidence exists of the use of these registries to aid care coordination.

2.1.7.6 Shared Repositories

As stated earlier, healthcare delivery through the care continuum require specific information essential for patient care across multiple institutions at different times. In practice,

there is consensus on the value of sharing health information for improved care. However, *'what needs to be shared'* either changes with the need or remains a source of constant debate. One approach is to adopt a *use case*, where implementers agree on a *use case* and determine how to share information amongst stakeholders for the *use case* e.g., Maternal health. The different parties then agree to the datasets to be shared.

In this thesis, however, the International Patient Summary (IPS) is used. The IPS, for instance, is an international movement to standardize cross-jurisdictional minimum data sets for Patient information sharing. The IPS is based on CEN EN 17269 and ISO/DIS 27269 and intended for specialty agnostic, condition independent, and unplanned care across national borders [57]. The IPS FHIR bundle contains a number of FHIR resources grouped for sharing Patients' health information [57]. The IPS(v1.0.0 Standard for Trial Use (STU)1) based on FHIR R4 require that the bundle at a minimum contain three resources as follows:

- Medication summary - Medication statement resource or Medication resource
- Allergies and Intolerances - Allergy Intolerance resource
- Problem list - Condition resource

The FHIR IPS bundle for exchange can contain other recommended or optional resources as necessary. The following FHIR resources are recommended: Immunization, Procedure, Organization, Performer, Observer, Device, Device Use Statement, Observation, Media observation, DiagnosticReport, Specimen, Imaging study, and Practitioner. Vital signs, Care plan, consent, and clinical impression FHIR resource are optional.

2.1.7.7 Privacy and Security in healthcare

Privacy standards help ensure a consistent strategy to protect an individual's (or organization's) right to "what is collected about them", "what is shared about them", "when it is shared", "who can access and used", and "for what purpose". The leading global standards (and regulations) for data privacy are the Health Insurance Portability and Accountability Act (HIPAA) [65] and the General Data Protection Regulation (GDPR) [66].

2.1.7.8 Healthcare Revenue, Incentives, and Token

There are two broad healthcare payer models. One is the government payer system like in the UK and EU, and the second is citizen payer (either out of pocket or through insur-

ance), as is the case in the US. Nevertheless, there are combinations between these two extremes. Point-based system for incentives in healthcare is gaining traction, from *mobile conditional cash transfer* to patients in Nigeria [67] to *meaningful use* (now *promoting interoperability program*) payment for service providers in the US [68], its value has been long established. Incentives aim to drive healthcare uptake and service provision behavior changes. Healthcare insurance systems use point-based (or tokens) systems for revenue computation. These concepts, when digitally implemented, are invaluable irrespective of the HIE model. Healthcare incentives and tokens are currently marginally implemented by HIE intermediaries. An emerging and popular kind of token is based on blockchain. Blockchain tokens can be minted, transferred, or owned for a stake. Blockchain based-token is gaining popularity because participants in a shared network can have visibility in how tokens are minted, managed, transferred, or owned. Despite the value of blockchain-facilitated tokens, there is limited evidence in the literature of their use in healthcare.

2.1.7.9 Public health aggregation & Epidemiology

Policy and public health decision-making often rely on aggregate health information. In addition, epidemiology disease surveillance depends on these aggregated summaries, as the present pandemic has shown. While this is not the area of focus of this research, aggregate summary data remain important for these reasons.

2.2 Health Information Exchange (HIE) state of the art

Akhlaq et al. already evaluated the evolution of Health Information Exchange (HIE) and its definition over time [69]. In this study, a systematic search of scholarly literature for evidence of HIE in healthcare implementation was conducted. First, discussing global HIE perspectives looking at European Union (EU)&United Kingdom (UK), United States (US)&Canada, and Low and Middle Income Countries (LMIC)s. The barriers to data sharing, the uses, and the effectiveness of HIE was discussed. The different HIE architectures and interoperability components were reviewed.

2.2.1 HIE initiatives

Table 2.1 illustrates the categorization of Health Information Exchange (HIE) systems by their geographic distribution and architectures used.

Table 2.1: Classification of Health Information Exchange strategies

Classification	Reference
Geographic distribution	*Africa-* [71] [72] [73] [74] [75] [76] [48] *UK & EU-* [6] [70] [77] [77] [59] [78] [79] *US & Canada-* [80] [81] [10] [78] [47] [82] [83] [84] [85] [69]
Barriers	[86] [87] [88] [89] [90]
Uses & Effectiveness	[91] [92] [93] [94] [95] [96] [97]
Standards used	[98] [99] [57] [58] [100] [101] [102] [13] [60] [12] [15] [103] [58] [104] [105]
Architecture	*Central-* [106] [107] [108] [10] [82] [83] [71] [109] [72] [73] [59] [110] *Federated-* [72] [111] [10] [47] [112] *Patient-controlled-* [50] [113] [114] [115] [116] [117] [118] *Decentralized-* [119][120][121][122][123][124][125][126][127] [128] [129] [130] [131] [132] [133] [134] [135] [136] [137] [138] [118] [139] [140] [141] [142] [143] [144] [1]

EU and UK HIE initiatives

Denmark has linked primary care doctors with laboratories, pharmacies, and hospitals through the MedCom initiative established in 1994. The use case started with referrals, discharge letters, laboratory reports, radiology, prescriptions, and claims [6]. The exchange was based on the European TC251 standard by European Committee for Standardization (CEN). By 2002, the Danish efforts had resulted in a 70 percent reduction in errors and significant information sharing amongst stakeholders [70].

Estonia's successful nationwide EHR implementation is facilitated by their famous X-Road, an interoperability layer between cross-sector services, including Banking, Education, and Healthcare [77]. The X-Road interoperability layer, launched in 2001, grew over the years from the first use case. The e-Health service component was added to X-Road in 2008. As of 2018, over 99% of medical prescriptions were digitized. Estonia's X-Road, as of 2020 connects over 2700 multi-sectoral services spanning 700 organizations (institutions and enterprises), including healthcare [77]. The Open Electronic Health Records (OpenEHR) is widely used for health data specification, clinical modeling, and software customization. OpenEHR has active communities in Japan, China, Germany, Brazil, and Spain [59].

In Europe, efforts to catalyze interoperability have been ongoing. Notably the EU EHR exchange format was facilitated by the EU 's eHealth Network [78]. This leading effort

for HIE in EU is codenamed e-Health Digital Service Infrastructure (eHDSI), initially aimed at cross-border information sharing for two use cases: patient summaries and ePrescriptions. The pilot initiative that facilitated HIE between Estonia and Finland only happened recently in 2019 [78]. The target is to have 22 EU countries exchange these two use cases' information by the end of 2021 [78]. There is limited information on the evaluation of eHDSI, which may be because it is still in the early stages of such evaluation. The eHDSI 's target is to leverage the existing European Reference Networks (ERNs) membered of about 900 specialized healthcare units located in 300 hospitals in 25 EU member states. The Nordic Interoperability Project also aims to connect five Nordic countries (Denmark, Finland, Iceland, Norway, and Sweden). Others mentioned are those by the Portuguese Ministry of Health, Shared services department [78]. The Intersystem HealthShare HIE solution describes support for Enterprise Master Patient Index (EMPI) that acts as a single source of truth for identifying patients in the system. No information about the coverage and impact. An InterSystems 2019 report indicated that their HealthShare HIE in the Netherlands had received consent from 13 million of 17 million residents to share their health information [79]. Also, HIE in the Netherlands are as follows: ZorgNetOos, Jeroen Bosch Ziekenhuis, Limburg Exchange Network, Image Exchange South East Brabant, Regional Exchange Network West Brabant, Image Exchange Network Breda, Rotterdam Exchange Network RijnmondNet, TRIJN, Zorgring Noord-Holland Noord. However, little information exists of their reach and impact.

2.2.1.1 United States (US) and Canada HIE initiatives

The US HIE initiatives were the center point of the 2009 Health Information Technology for Economic and Clinical Health (HITECH) Act in the US. However, a decade after, HIE adoption remains below key targets [80]. The Regional Health Information Organization (RHIO)s are popular in the US's HIE space. Healthix is one such RHIOs indicating on their website to have data of over 20 million patients from more than 8,000 health facilities representing 2,200 healthcare organizations [81]. Healthix covers health facilities in New York City and Long Island in the US. A July 8, 2021, blog post on their website titled "Healthix VIPs Deliver Healthcare Interoperability Solutions" launched a Vendor Interoperability Program (VIP) to support vendors to interoperate on the Healthix network.

The Indianan Health Information Exchange (IHIE) was originally started as a community information exchange initiative by Regenstrief Institute to connect cross-county practices by Indiana School of Medicine practitioners and Veterans [10]. The IHIE went through several evolutions to the point where a non-profit was registered to centrally manage

the information exchange initiative. IHIE non-profit had board of directors from 16 organizations that included *Hospital networks*, *Public agencies*, *Medical societies*, *Physicians*, *Consumers*, and *Researchers* [10]. IHIE has now grown to include 15 billion clinical data elements contributed from 18 million unique Patients, 53,004 Practitioners from 18,738 health practices, and 153 hospitals [82]. The IHIE contained patient information segregated by the source institutions and they include: Emergency visit data, demographic records (registration), Radiology report, Discharge summary, Medication summary, EKG report, Laboratory data, Encounter data, coded diagnosis and procedures, and Ambulatory encounter [82].

Another popular centralized model is the New York HEALTHeLINK HIE initiative. HEALTHeLINK, initially established by a network of payers and 26 hospitals to support administrative data exchange, was later expanded using a public grant to cover HIV and tuberculosis (TB) [47]. HEALTHeLINK website did not show public implementation details of data sharing.

The HealthEConnections brings together patient medical information scattered in hospitals, practices, laboratories, and imaging centers in a State Health Information Network of New York (SHIN-NY) [83]. HealthEConnections is a qualified entity of SHIN-NY that governs the flow of patient data among healthcare providers. HealthEConnections indicated on their website that in 2018 alone, 7.5 million alerts were delivered from 600,000 unique patient information [83]. They noted that 2700 physicians from 860 organizations participate in the exchange [83].

The government of Minnesota in 2007 passed legislation that required "all healthcare providers in the state to implement an interoperable Electronic Health Records (EHR) system by January 1, 2015 (Minn, Stat. §62J.495) [78]. The rationale for the mandate was that it would help with more effective use of health information through the timely exchange of such information. In 2019, the Minnesota e-Health Advisory Committee and the Minnesota department of health recommended that all providers only demonstrate progress toward interoperability. In June of 2019, the electronic health record mandate was eliminated [78]. Also, the Midwest health connection is indicated on its website as having a database of over 28 million electronic patient records from the Midwest US. The network uses a network of participating members and selects regional gateways to connect to the central repository. Details were not provided about the architecture and modus operandi on their website [84]. HealthiE Nevada is another private community-based HIE network established to serve healthcare stakeholders in Nevada state US. On

their website, a list of participating stakeholders, including HIE organizations and hospitals, is provided [85].

2.2.1.2 Africa HIE initiatives

There was limited evidence of multi-stakeholder HIE implementation in any African country. However, active development appears to be ongoing as African institutions have published a few proposed HIE frameworks. Angula et al. developed a prototype for aggregating data from regions in Namibia to a central District Health Information Systems version 2 (DHIS2) Health Information System (HIS) repository [71]. An integration of health information systems in Rwanda using OpenHIE was equally proposed [108]. Similarly, a result from the systematic survey of technical aspects of HIE, James Gor, in 2017 proposed an agent-based Health Insurance information exchange in Kenya [72]. During the same period, Brenas et al. presented an agent-based malaria sentinel sites integration in Africa with Uganda and Gabon as proposed test cases [145]. However, from the search five years after (2021), little evidence exists of progress beyond these proposals. Chris et al. indicated using Open Health Information Mediator (OpenHIM) in South Africa to connect SMS message aggregate to mothers with the DHIS2 platform [75]. For this use case, it was not clear if the information exchanged involved more than one organization.

As the scholarly search yielded limited evidence of HIE in Africa, it was augmented with a traditional google search. On the OpenHIE, website, several case studies were listed [48]. While there are case studies on their website for Tanzania, Zambia, Malawi, Rwanda, South Africa, and Ethiopia, little information exists on actual multi-organization information exchange. Another result from the open-source LMIC search was the agreement by the Benin Republic and the eGovernance academy in Estonia in 2018. The agreement is intended to facilitate the design of the X-Road-based interoperability layer using the digital ID as the first use case. However, no recent detail exists of the progress made. Other results show that Ethiopia piloted the use of Open Concept Lab (OCL) system for terminology information management. The Integrated Human Resources Information System (iHRIS) software, used for health worker information management, has been piloted in Botswana, the Democratic Republic of Congo, Ghana, Guinea, Kenya, Lesotho, Liberia, Malawi, Mali, Nigeria, Rwanda, Sierra Leone, South Sudan, Tanzania, Togo, Uganda, and Zambia [76]. Similarly, the OpenCV for identity and civil registration is on trial in Zambia. Despite all ongoing initiatives, the search did not find any multi-organization HIE.

2.2.1.3 Asia HIE initiatives

Similarly, Khalique et al. developed a framework for data acquisition, standardization, and reporting in Pakistan, another typical LMIC [73]. Also, a recent healthcare stakeholder survey in Pakistan found no effect (or evidence) of HIE implementation in the country [74]. Some reasons for the lack of data sharing progress include health providers' concerns about liability consideration and poor documentation skills.

2.2.1.4 Other HIE initiatives

Other global HIE initiatives as itemized by [69]: Accenx (US), 4Med(US), Alaska eHealth Network (AeHN) (US), Maryland's HIE (US), Alert HIE (Portugal), Arkansas State Health Alliance for Records Exchange (SHARE) (US), NeHII Nebraska (US), Centricity (US), Cerner (US), Pennsylvania ClinicalConnect (US), Codagnon (EU), CORHIO (US), CSC (Australia), Emdeon (US), Excelicare (UK), Florida HIE (US), Florida HIN (US), Forcare (Netherlands), Georgia HIV HIE (US), Health insights (US), Health Unity (US), HIE Bridge (US), HIE Ohio (US), HINAz (US), Illinois HIE (ILHIE) (US), Inteli Chart, North Dakota HIN, Oracle (US), Orion HIE (US), SouthEast Michigan HIE (US), Texas HIE (US), Utah state Clinical HIE (cHIE) (US), and Xerox HIE (US).

2.2.2 Barriers to healthcare data sharing

Health Information Exchange (HIE) in practice remains challenging. Panguise et al. conducted a systematic literature search and review of the potential barriers to clinical and public health information sharing [86]. The authors found six main barriers that limit clinical and public health information, as listed. The six barriers identified include: Technical, Motivational, Economic, Political, Legal, and Ethical. Each of these barriers contribute in one way or another in limiting ability to share health information. Similarly, a recent study of inhibitors to HIE found fourteen inhibitors limiting clinics from sharing ambulatory information with other clinics in the US [87]. The leading technology inhibitors were inadequate technology infrastructure, difficulty integrating external data with EMR systems, and data security concerns. Clinic-to-clinic information sharing was equally limited by the inadequate capacity of the support personnel. The environmental inhibitors included "..legal concerns and complexity in framing HIE agreements with partners" [87].

In 2015, the government of Switzerland passed legislation mandating hospitals to adopt interoperable EHRs [88]. The reform's objective was to enhance data sharing amongst

providers to drive up quality healthcare and service efficiency. Despite the legislation, adoption was slow, and proposals for incorporating incentive mechanisms post-project-pilot were favorably recommended. Legal, privacy and ethical barriers have often been cited as key inhibitors to HIE adoption and use. However, Mello et al., in a US review of policy and legal frameworks, found that most of the legal barriers to HIE adoption are no longer reasons for concern [89]. Nevertheless, adoption remains sub-optimal despite HITECH facilitated incentives. A Korean implementation and documentation of the approach, including barriers encountered in the adoption of HIE, found poor architecture and standards design, documents and data issues, consent management, and usability as primary barriers [90].

2.2.3 Uses, effectiveness, and outcome of HIE

Hersh et al. also conducted a systematic review to determine the evidence of effectiveness, impact, and current uses of HIEs [91]. They found varying evidence of HIE in all 34 locations in their study design. Eight were in the US, and there was no study in Africa. Also, they found only Finland and Austria in the EU block. Similarly, Dobrow et al. conducted a systematic review to ascertain the international evidence base of adoption and use of EHRs [92]. Their measured health outcome was overall positive though highlighting the poor study design of most research.

HIE is used for Ambulatory laboratory testing and claims reporting [93]. They found the number of tests in the claims submitted by 34,818 patients by 306 practitioners from 69 practice institutions increased after HIE adoption [93]. In another study, the use of HIE for referral in emergency departments found improvements in care quality, time savings, and cost savings [94].

Tzeel et al. found that Patients saved as much as $29 per emergency department visit when they used HIE [96]. They also found that in general, there was a decreased utilization of imaging procedures and diagnostic tests. Frisse investigated the financial impact of HIE in emergency care in all major emergency departments in Memphis and Tenessee [95]. They studied 15,798 encounters for financial impact and found that a reduction in admissions and laboratory tests resulted in $1.9million in annual savings.

Walker D.M. also sampled 1017 hospitals to determine the effect of HIE participation of hospitals on efficiency. The survey found that HIE participation by hospitals can improve technical efficiency and Total Factor Productivity (TFP) [97]. Dixon et al. evaluated 7.5

million laboratory reports in the US and found that data of patients processed by HIE were more complete than others. Their findings suggest that HIE can help improve data completeness, reporting, and quality. Vest and Miller and their study of 3,278 hospitals found improvements in patient satisfaction scores in hospitals that adopted and implemented HIE.

2.2.4 Digital health standards

As noted in chapter one, standards in healthcare can e grouped in many different ways. A popular grouping covers data transport, content-structure, vocabulary (terminology), identifier, privacy, safety, and security. Communication standards include CEN EN 12052: 2017 [98] - Digital Imaging and Communication in Medicine (DICOM) for image communication, workflow, and data management. The healthcare document interchange format, Hl7 v2 and HL7v3 [99]. The HL7 FHIR v4 [57] uses a RESTful lightweight health resource exchange using API. Others are the CEN/TC 251 - ISO 13606-1 -5:2019 [58], CEN EN 1064: 2020 [100], CEN ENV 13607:2020 [101], ISO CD TS 22691 [102].

The National Health Management Information System (NHMIS) indicators is an aggregate standard . For terminology classifications, Logical Observations Identifiers Names and Codes (LOINC) [13], International Classification of Diseases (ICD) [60], Systematized Nomenclature of Medicines Clinical Terms (SNOMED CT) [12], RxNorm [15]. The privacy, security, and safety standards are: ISO DIS 27799 [103], CEN TR 15300, CEN/TC 251 - ISO 13606-4:2019 [58], ISO IEC 13888, ISO TC 215 -TR 21098:2018 [146], ISO/IEC TR 14516:2002 [147], ISO IS 17090-1: 2013 [148], ISO IEC 10118-3: 2018 [149], CEN CR 13694 [150], CEN TR 15299 [104], ISO TR 21730:2007 [151], CEN EN 14485:2004 [105]. When specifying requirements, the standards CEN/ISO EN 13606:2019 [58] and ISO TS 18308:2011 [152] are popular.

2.2.5 HIE architectures in literature

HIE systems can be differently classed by the data and service governance control, ownership, or data flow direction. As noted earlier, when classed by governance, a HIE system can be either Centralized, Federated, Patient-Centered, or Decentralized. A particular HIE can take combined features of more than one of these architectures. When HIE is classed by ownership, they can be Private, Government-facilitated, or Community-based HIE networks [111]. When viewed by data flow direction, an HIE can be Directed, Query, and Multiple or combinations thereof. In July 2021, a systematic search and review of

literature in IEEEXplore to ascertain technical aspects of HIE in the last five years (2017-2021) was conducted. The subsections that follow discuss the findings, amongst others, grouped according to the architecture proposed or used.

Centralized HIE models

Most prototypes and implementations of HIE architectures use centralized architectures either on centralized hardware or distributed hardware managed by one actor. An actor here is relative to the organization, not ownership of such organizations.

Osei-Tutu et al. proposed a cloud-based centralized framework for HIE and illustrated them using the Zachman and The Open Group Architecture Framework (TOGAF) Frameworks for presenting the enterprise architectures [106].

Similarly, the OpenHIE architecture is modeled as a centralized architecture with service, point-of-care, and interoperability layers [107]. The OpenHIE interoperability layer leverages the SOA for implementing the central interconnectedness of connected systems. A Software Defined Network (SDN) was proposed for centralized integration of health information systems in Rwanda using OpenHIE [108]. OpenHIM promoted by the OpenHIE community and discussed in the preceding chapter use the centralized architecture. At the time of finalizing this search was conducted in September 2021, these initiatives are either at the proof-of-concept or pilot stages.
Similarly, the Indiana IHIE and the New York HEALTHeLINK described in the preceding section are also forms of centralized HIE models [10] [82]. Organizations in the HealthE-Connections's New York State information network connect through *myConnections*, a secure, single sign-on portal for authorized users rely on this centralized *myConnections* platform [83].

In Namibia, the proposal for centralized data aggregation was prototyped [71]. Yang et al. proposed Medshare for Patient-centered care. However, the proposal uses a centralized cloud [109]. In Kenya, it was proposed to connect health insurance systems using a centralized architecture [72]. The architecture proposed centralizing tasks and distributing resource storage. Also, the model by Khalique et al. in Pakistan is a full-fledged centralized HIE model [73]. The OpenEHR is also a centralized architecture for information sharing [59]. Though no result of technical aspects of implementing OpenEHR was returned in the IEEEXplore systematic technical information search [118][153][110].

2.2.5.1 Patient-centered HIE model

In a Patient-centered HIE model, the Patient is in control of their data either directly or through a technology service provider. Proposals and prototypes for Patient-centered care are widely published [50, 113, 114, 115] although using blockchain to facilitate this organization-level exchange. Vest and Miller proposed a hybrid centralized facilitated Patient-centered HIE system [116]. Yan et al. discussed directed, query-based, and consumer mediated HIE architecture for Patient-centered information exchange using blockchain [117]. OmniPHR, a patient-centric information sharing framework and platform gives a patient control of their PHR, leveraging blockchain. At the same time, the service provider and other relevant stakeholders also have access to the patient EHR, hitherto scattered in different custodian health institutions [118].

2.2.5.2 Federated HIE model

A Federated HIE architecture is such that participating organizations keep control and storage of their generated data, and others request to access the data on demand. Very few systems have been identified in the literature as federated. In Kenya, it was proposed to connect health insurance systems using a centralized architecture [72]. The architecture proposes centralizing tasks and distributing resource storage. Dixon noted that Federated architecture had been successfully used in Indiana Health Information Exchange and the New York City Primary Care Information Project [111]. Though, my August 2021 search did not return results of these successful ac HIE integrations. Though, it is possible that Indiana's IHIE [10] and New York's HEALTHeLINK [47] both use a combination of centralized and federated architecture to varying degrees. The X-Road which the Estonia's HIE is based, has been described as having the features of a federated system where each service provider generates and shares data as appropriate, only that they have to be online [112].

2.2.5.3 Decentralized HIE model

In a decentralized system, the generated data or the processing work does not have to happen on the generating organization's system. At the same time, the data is distributed across multiple participating actors in information sharing. Blockchain is emerging as a model of choice for trustless information sharing, particularly in healthcare. As the application of blockchain in healthcare is central to this thesis, an in-depth discussion is provided in the next section.

2.3 Blockchain in healthcare

A comprehensive systematic literature review of how blockchain is used in healthcare (including for HIE) was conducted by this author in December 2019 and published [1]. Based on the search and analysis, prototype and pilot papers that discuss data sharing, storage, or data processing using blockchain are shown in Table 2.1.

2.3.1 Distribution by Function

Leveraging the functional distribution outline by Hölbl et al, the articles were analyzed for trends [154]. Please see (SeeFigure 2.6). The use case of blockchain in healthcare that received most attention are the data sharing use case and the access control use cases.

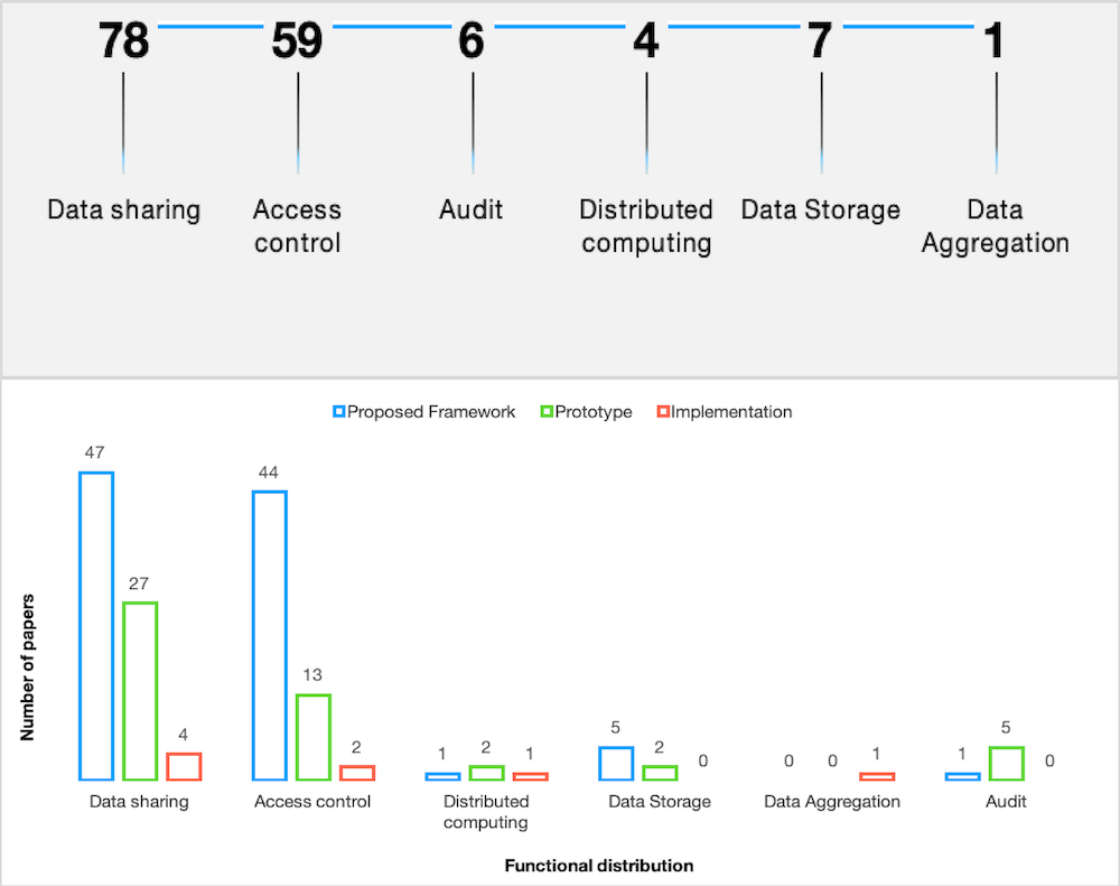


Figure 2.6: Functional distribution of articles [1]

Access control with a focus on identity management like was a specific use cases. Nonethe-

less, a handful of the articles discussed auditability, non-repudiation, data storage, distributed computing, and data aggregation. The analysis of published literature show that data sharing and access control represent majority of papers by functional distribution. Some papers proposing use of blockchain in healthcare considered the following service delivery use cases across multiple healthcare sub-domains, which we already detailed in our published paper [1]. The domain varied: Arrhythmia image classification, blood-management, Cancer care, clinical trials, Diabetes, Dental care, Dyslexia care, DNA compression, Dermatology Radiology, Insurance, Oncology, Provider communication, Supply chain, and Universal Health Coverage (UHC). A national-level health system application was discussed in one implementation, though there were limited details. See Table 2.2 for additional details of the paper analysis largely adapted from our already published and highly cited work [1].

2.3.2 Privacy and Security analysis

The articles were further analyzed for contribution to privacy and security state of the art in health information while complying with existing regulations. Like most Information Technology (IT) systems, a blockchain system faces many risks. Notably, private key security, Double spending (which was not a problem for traditional centralized IT systems), Criminal smart contracts (trying to manipulate blockchain logic for criminal gains), Byzantine fault attempting to execute unauthorized transactions or chain take-overs), 51% attack vulnerability, vulnerable smart contracts, under prices transactions, over priced transactions [248]. Other papers did not provide adequate detail for security analysis, and such papers were excluded from the review. Also, detailed security analysis was beyond the scope of our current thesis.

There was limited discussion on established privacy and security regulations like Health Insurance Portability Accountability (HIPAA). A few articles noted compliance with Health Insurance Portability and Accountability Act (HIPAA) [157] [165] [188] [190] [191] [120] [123] [125] [118] [239] [199] and with the European General Data Protection Regulation (GDPR) [166] [239]. Kristen et al. reported compliance the HIPAA guideline with respect to Protected Health Information (PHI) [123]. In the definition of HIPAA, the PHI is a set of information datasets that, if revealed either individually or in combination with other data, has the potential to lead to the ability to uniquely identify the owner [123]. Examples of such datasets can be names, phone number, and address. The US's Health Insurance Portability and Accountability Act (HIPAA) provides a complete list for reference and updates the list regularly.

Table 2.2: Classification of solutions by frameworks, prototypes & implementations [1]

Classification	Reference
Conceptual Models, Frameworks or Proposals	*Data sharing*- [155] [156] [157] [158][159] [160] [161][162] [163] [164] [165] [166] [167][168] [169] [170] [171] [114] [172][173] [174] [175] [176] [177] [178] [179] [180] [181] [182] [183] [184] [185] [186] [187] [188] [189] [190] [191] [192] *Access Control*- [157]-[162] [164]-[166] [168]-[193] [173] [175] [177]-[179] [181]-[186] [188] [189] [190][191] [194] [195] [196] [197][198] [199] [110] [200] [201] [202] [203] [204][192] [205] [206] *Privacy*- [207] [208] [209] [210] [211] [212] *Audit*- [213] [198] [201] *Data storage*- [214] [215] [201] [203][204] *Distributed Computing*- [216] *Service delivery*- HIV[180], Clinical trial[156][160][186][217][218], Insurance [164][169], Diabetes[181][219], Cancer[190], Blood mgt.[220], Pharmaceuticals[162][221], Dermatology[171] LMIC[222], National Health System[223], Medical imaging[199][224][205], Health Education[202], Genetic data[225]
Prototypes and Experimentations	*Data sharing*- [119][120][121][122][123][124][125][126][127] [128] [129] [130] [131] [132] [142] [133] [134] [135] [136] [137] [138] [118] [139] [140] [141] *Access control*- [119] [120] [121] [123] [125] [226] [132] [138][118] [227] [135] [228][229] [230] [231] [232] [206] [233] [234] *Privacy*- [235] [236] [237] [217] [238] *Audit*- [119] [120] [122] [125] [134] [239] *Distributed computing*- [142] [240] *Data storage*- [128] [241] *Service delivery*- SoftwareDesign[135] [242] RadiologyOncology[139] [126][139], Cancer [139], SupplyChain [127][140], PatientCenteredCare [122] [128] [131] [118], ProviderCommunication [243], DNA Compression [142], ArrhythmiaImageClassification [227], HaemoglobinTest (HbA1c) [239], RemoteCare [244]DentalCare [211]
Implementations or Pilots	*DataSharing*- [142] [142] [143] [144] *AccessControl*- [143] [245] *DistributedComputing*- [142] *DataAggregationAnalysis*- [246] *Service delivery*- Dyslexia [246], National Health System[223] SupplyChainManagement [247] [144]

The authors, Xia et al. in their work used cryptographic approaches and special database to verify cloud-cloud communication for blockchain-based health information-sharing [243]. Similarly, Alex et al., in their OmniPHR framework, used a configurable access control framework to give control to manage permissions themselves [118]. They simulated the permissioned-blockchain access management and the encryption algorithm used was not discussed.

The proposals by Alevtina et al. show a patient's ability to read, write, or share, linked unique data access to the linked provider institution [139]. The Hyperledger fabric chaincode (equivalent of traditional smart contracts in Ethereum blockchain) was used to simulate management of all this. Similarly, Benchoufi et al. designed a framework proof-of-concept they used to enroll volunteers [141]. Volunteers could sign transactions on the web platform using their Public Key Infrastructure (PKI) generated key [61]. Similarly, Castaldo et all in their OpenNCP model, discussed their inter-jurisdictional HIE platform highlighting auditability, patient ID, and notification functionalities [61] [122]. OpenNCP is designed to annually regenerate keys and share with internal Certificate Authority (CA)s. Other research emerging in this domain is the keyless signature mechanism used for access control and management [135].

Ji et al. also proposed "BMPLS: Blockchain-Based Multi-level Privacy-Preserving Location Sharing" [249]. The BMPLS is a proof-of-work based, privacy-preserving, and order-preserving framework for encrypted patient geo-enabled location information sharing. This framework used the RSA-1024, OPE(2), and SHA-256 algorithms. Similarly, Mikula et al. in Denmark, used a blockchain-based access management framework to evaluated and show that authentication and authorization for all 3.8MB provider data took 3 seconds [229]. The Ethereum network was used for the implementation.

Estonia is shinning light for Health Information Exchange (HIE) success story. This is because 99% of patients in Estonia has access to their digital health record in one form or another (directly or indirectly) [250][251]. Also, an estimated average of 300,000 patient queries happen annually. Electronic healthcare billing information is conducted 100% electronically. Estonia is the first national level blockchain system constituting of three decentralized blockchain networks. The first is a permissioned public and shared blockchain based on Ethereum and Bitcoin, the second is a permissioned private blockchain based on Microsoft Coco, and the third is a permissioned Hyperledger Keyless Signature Infrastructure (KSI) [252]. The organization Guardtime is responsible for managing the KSI, whose

main goal remain facilitating regulatory compliance, integrity assurance, and dimensioning. No information currently exist to show that this system has been evaluated. This program popularly called X-road is the only claim of national level use of blockchain in healthcare. Though details are not public to fully understand the operations outside of what these blogs and articles provide. Also, in other gray literature, the X-Road has not been presented as a blockchain solution.

2.3.3 Performance analysis

In the prior published paper, we assessed the Prototypes and Implementation (or Pilot) papers for performance evaluation using the Hyperledger Performance evaluation metrics with the result as in Table 2.3 [253] [61]. Each paper was evaluated to determine whether a simulation was performed and the outcome. The main elements checked for each paper were the number of observations, the type of tool used, the duration of the test, transaction throughput, and the type of hardware used. Simulations were mostly conducted using a laptop PC with test tools ranging from JMeter, Geth Sim, Remix, and many others. The number of nodes also varied compared to the number of observations. It was difficult to ascertain the size of resources used per transaction in a uniform way, for instance, information on Random Access Memory (RAM) utilization, Central Processing Unit (CPU), and disk-space utilization for each type of transaction. In this study, the RAM, CPU, network, and disk-storage utilization for each blockchain operation have been captured from the simulation.

2.3.4 Blockchain-facilitated Identity

The search and analysis found broadly three types of identity management for permissioned blockchains using Certificate Authority (CA). A CA is the infrastructure that creates (generates), stores, manages, distributes, and revokes identities and signature keys. The three possible architectures as in Figure 2.7 emerged, which has now been published [61]:

- *One CA blockchain*,
- *Multi CA blockchain*,
- *Client Self CA blockchain*

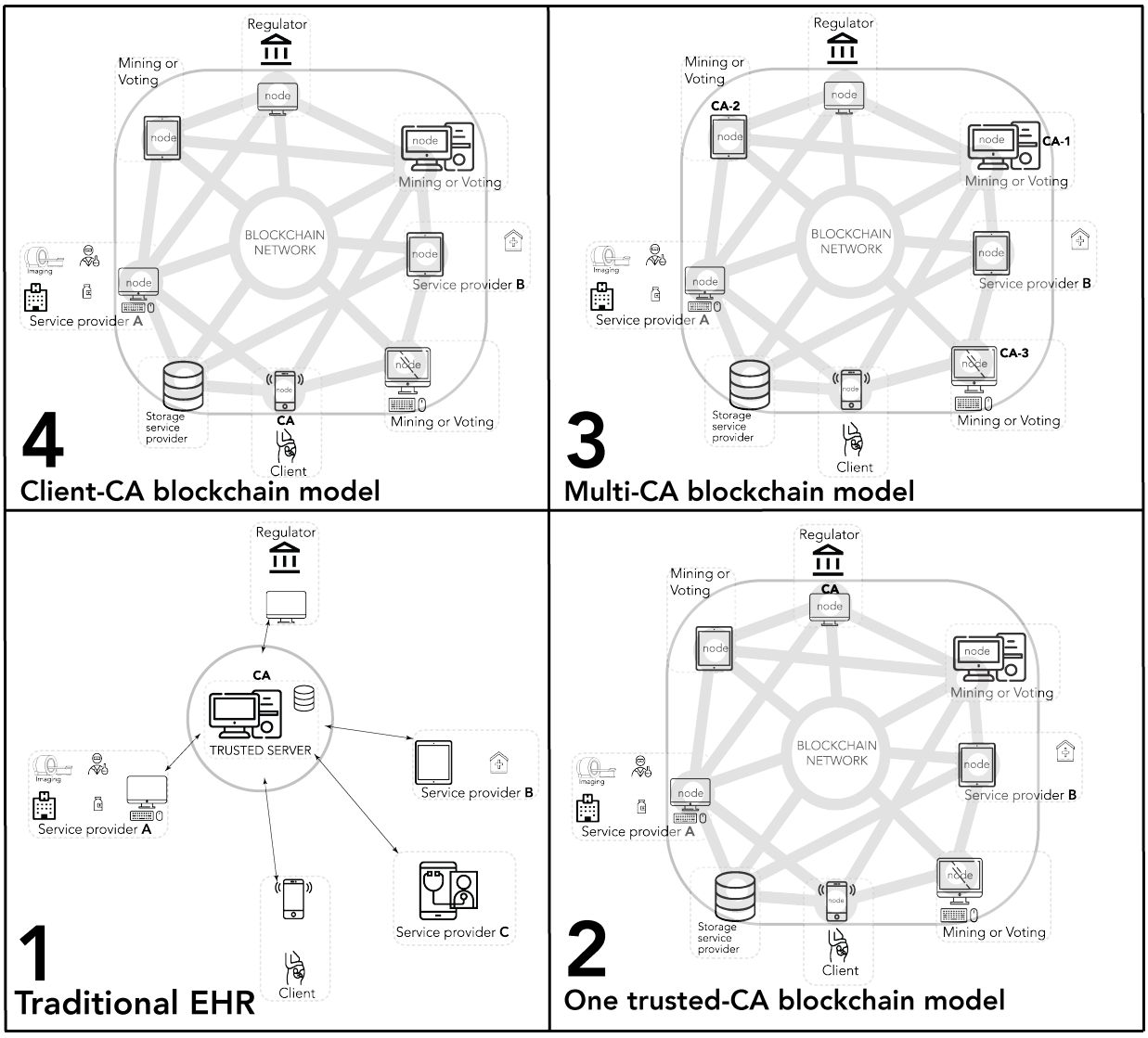


Figure 2.7: Blockchain in healthcare CA management strategies [1]

Table 2.3: Comparing blockchain in healthcare prototypes and pilots [1]

Article	Observations	TestType	TestCount	Duration	Tx Latency	Tx Throughput	HardwareSpecs
[243]	5-100 users	JMeter sim.	8 tests	54-1286 secs	NA	NA	NA
[118]	100-3200 nodes	OverSIM sim.	20 tests	3 hours	500ms	NA	2GHz, 2-core, 8GB
[247]	IoT sensors	Pilot study	7576 points	1 month	NA	NA	NA
[121]	4 VMs	Experimental	1 test	5 days	NA	NA	i5, 2.2GHz, 8GB
[128]	8 servers/clients	Experimental	56 Tx	NA	NA	46 Tx/s	NA
[254]	1 - 50MB file	Geth sim.	50-350 Tx	N/A	200-500ms	30MB file	i5,2.2GHz,16GB,Mac
[129]	1 node	NA	NA	N/A	0-5 secs	NA	i5, 2.5GHz, 8GB
[130]	NA	NA	5-100 con.	NA	0-900ms	NA	NA
[132]	10 nodes	Experimental	1-10 nodes	NA	23 - 205 secs	NA	i7,2.7GHz,4GB
[153]	64-512 con.	Pilot study	40,000pers	1wk	184-556ms	26-278MB	NA
[142]	NA	NA	25 samples	NA	NA	NA	NA
[228]	2 nodes	MIRCL cpabe sim.	10 runs	NA	NA	NA	i5, 3.3GHz,8GB
[133]	NA	Virtual 3D Avatar	NA	NA	NA	NA	NA
[255]	2-4 communities	Experimental	NA	NA	5-30mins	NA	i5, 3.3GHz,8GB
[229]	1 PC	Chrome sim.	4000 blks	NA	219ms	767kB	2.4GHz, 8GB,1TB
[135]	1 PC	JMeter sim.	NA	NA	500ms	2048kB	NA
[256]	N/A	Pilot study	15 pers	NA	14 secs	NA	NA
[223]	N/A	Implementation	NA	NA	Ongoing	NA	NA
[257]	1 node	Java-pair cryp.Lib	NA	NA	1.5ms	NA	i7,2GHz, 8GB, Win10
[244]	multiple	Remix	NA	NA	NA	NA	NA
[144]	1 node	AWS cloud IoT	25tx	NA	NA	NA	AWS EC2 instance
[258]	2 node	AWS cloud	500 pxts	NA	<=30sec	8 - 128kb	ES2, Ubuntu16.04
[259]	2 node	AWS cloud	10patients	NA	<=40ses	50 - 300kb	ES2, Ubuntu16.04
[231]	2 node	RaspberriPi	2-10 blks	NA	200ms-2sec	8 -128kb	Rasp.Pi 3.3GHz
[236]	1 node	AWS cloud	NA	NA	5-6sec	236bytes	ES2, Ubuntu16.04
[260]	1 node	Remix	NA	NA	NA	NA	NA
[237]	1 node	Experimental	NA	NA	7-12 Mb/ms	21-36kb	NA
[205]	1 node	AWS cloud	50 users	9days	NA	259.2kb	NA
[261]	1 PC	NA	NA	NA	NA	NA	ES2, UbuntuLT16.04
[217]	10 PCs	NA	NA	NA	NA	NA	Win10, i7,3.4GHz
[233]	4 PCs	Experiment	NA	NA	NA	NA	Xeon,ES2,Ubuntu

MedRec [120] and Medblock [142] both follow the architecture of a centralized, One-CA where the Identity used is managed centrally. On the other hand, the OpenNCP architecture proposes one where nodes generate and manage their keys [122]. A trusted CA is responsible for validating the key pairs. On the other hand, the OmniPHR model proposed a client-CA architecture as we have shown in Quadrant 4 provides a mechanism allowing the client full control of the identity and PHR [118]. Though article indicated that this architecture of OmniPHR is configurable to a Quadrant 2 model to act as trusted CA organization. An example of a Multi-CA architecture was proposed by the CB-SIFT framework for encrypted health information extraction using blockchain [255] . The CB-SIFT framework was benchmarked against the bitcoin, and show better performance over bitcoin. Yet, bitcoin is hardly the best benchmark for evaluating optimal performance of blockchain network as it has notoriously poor transaction throughput because of its use of proof-of-work consensus algorithm. While MedShare was evaluated, it provided little detail on how user authentication happen, which can either mean that the framework is either using a Trusted CA architecture or a Multi CA architecture [109]. Also, another paper uses a Hyperledger blockchain membership service that acts as a Certificate Authority (CA) to generate key pairs on the network [139]. Also, an attempted design of a national Master Patient Index (MPI) using blockchain was presented with little further detail [197] . In Estonia, the National ID is the unique identification scheme using a MPI model [250][251]. A centrally managed Patient identifier has to date been the recognized approach to Patient identification [62]. In practice, there are many other schemes used for Patient identification, of course with limited results - health institution ID, MPI, other functional IDs like drivers' license, Subscriber Identifier Module (SIM) and international passport [61].

2.3.5 Communication or Reference model standards

Here, the standards used by decentralized HIE systems based on the review is detailed. HL7 FHIR was by far the most used of all communication and information model standards [155][158][177][183][190][120][135][118][239][242] [238]. Older versions (v2 and V3) of HL7 were only used in [189][191][120][122][118]. See Figure 2.8 for the distribution of standards or ontologies mentioned in the reviewed solutions.

The OmniPHR model used super-peer node in a blockchain network that acts as a translator, and conditionally triggered depending on availability or not of passed data in the openEHR architecture format [118]. When triggered, the data is translated from one of the supported open standards to the OpenEHR. The system translates by

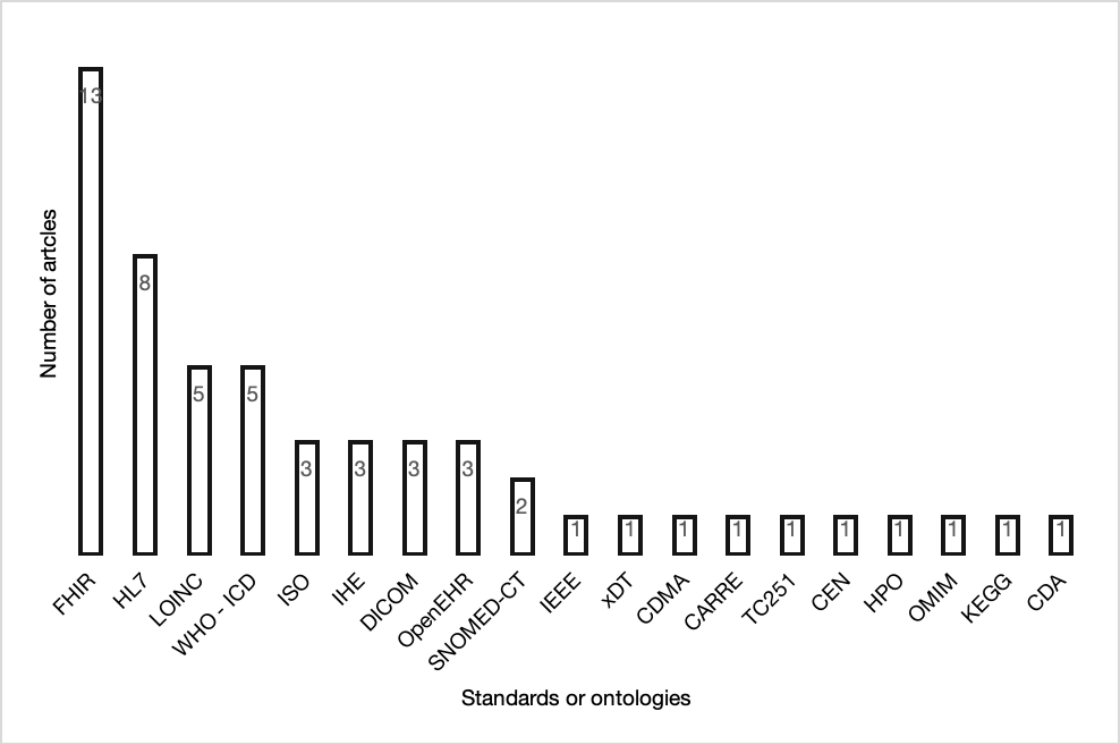


Figure 2.8: Healthcare ontologies and standards used [1]

- mapping open standards to OpenEHR standards.
- using Natural Language Processing (NLP) and Ontologies
- using Software agents mediation.

The only national blockchain initiative from my survey is the Estonia X-Road system that uses "..XML-based HL7 version3 messaging standard" [123].

2.3.6 Clinical terminologies (vocabularies) Registries

The Systematized Nomenclature of Medicines Clinical Terms (SNOMED CT), a poly-hierarchical health ontology with the largest number of clinical concepts, relationships, and detailed descriptions, is used [185][118]. It is poly-hierarchical because it supports multiple parents and multiple siblings. SNOMED CT contains 19 high-level clinical concepts -

- Body structure,
- clinical finding,

- environment or geographic location,
- event,
- observable entity,
- organism,
- pharmaceutical,
- biology product,
- physical force,
- physical object,
- procedure,
- qualifier value,
- record artifact,
- situation with explicit context,
- social context,
- special concept,
- specimen,
- substance,
- and SNOMED CT Model Concept

These high-level concepts are composed of sub-concepts based on relationships with parent or child concepts to yield an estimated 60,000 concepts [12]. Concepts have unique identifiers and a human-readable descriptions. The Logical Observations Identifiers Names and Codes (LOINC) though not as widely used in practice, is terminology for laboratory information and was used in proposals and prototypes for blockchain use in healthcare [158][179][191][118][171]. The Digital Imaging and Communication in Medicine (DICOM) used for standardizing clinical images by imaging equipment manufacturers is also used [191] [118] [199].

Classification systems are used mainly for reporting, do not show inter-relationships, and are uni-hierarchical. The WHO ICD versions 10 and 11 are the leading classifier schemes

used for concept coding in healthcare. Classification systems have one parent and one child. The ICD 10 database is estimated to have about 10,000 coded terms and is more widely used than SNOMED CT, especially in LMICs. The WHO's ICD was used by the proposed and prototyped systems [179][191][118] [230]. There was little detail on how the classifier was used in the studies.

2.3.7 Hyperledger in healthcare and HIE

Hyperledger Fabric is the leading enterprise platform for blockchain implementation that addressed the Proof-of-Work (PoW) energy consumption and sustainability problem of first-generation blockchains. Another systematic literature search was conducted in July 2021 for Hyperledger applications used in healthcare to ascertain the current state of the art related to my research on IEEEXplore. The choice of IEEEXplore was purposeful for identifying technical aspects of the published articles only. The search returned 56 articles, though only a few were of good quality and relevant to the HIE research. Singh et al. presented a detailed architecture and mathematical proof for Hyperledger use for in-hospital service tracking [51]. Their architecture was grouped into the following modules - Participants, Assets, Appointments, Transaction, and Chemists modules. They presented the performance using key Hyperledger fabric evaluation metrics. Others use Hyperledger for organ transplant accountability and waiting list management [262]; Insurance claims management using agreed data structure [263]; counterfeit drug management using RFIDs [264].

2.4 Other Healthcare Interoperability components

The background section of the introduction chapter itemized data structure standards, terminologies, identifiers, repositories, registries, and tokens as interoperability components of interest in this thesis. A deep dive and review of these components were done.

2.4.1 Traditional Identity management

Several strategies exist for Patient identity management, such as the use of the cross-sector national Identifier (ID) scheme, which is quite popular. Another approach is employing a central repository to manage a Patient's unique identifier, sometimes interchangeably called "Master Patient Index (MPI)" or "Client Registry (CR)". From my previous work, this could either be a "...national ID scheme, the health institution-managed scheme, an MPI, or other functional IDs" [1]. The publication indicated that effective Pa-

tients' ID scheme must consider patients enroll, authenticate, PHR security of storage, stored data governed, trust mechanisms adopted for decentralized governance is decentralized, duplicate management strategies, Identity creation and issuance strategies, and any other social determinants that may impact use [1]. Similarly, McFarlene et al. listed the characteristics that an ideal Patient ID needs to meet: They include: Uniqueness, Not changing, Ubiquitous, Not complicated, Not expensive, and Not controversial [265]. These competing characteristics have been illustrated in Figure 2.9 .

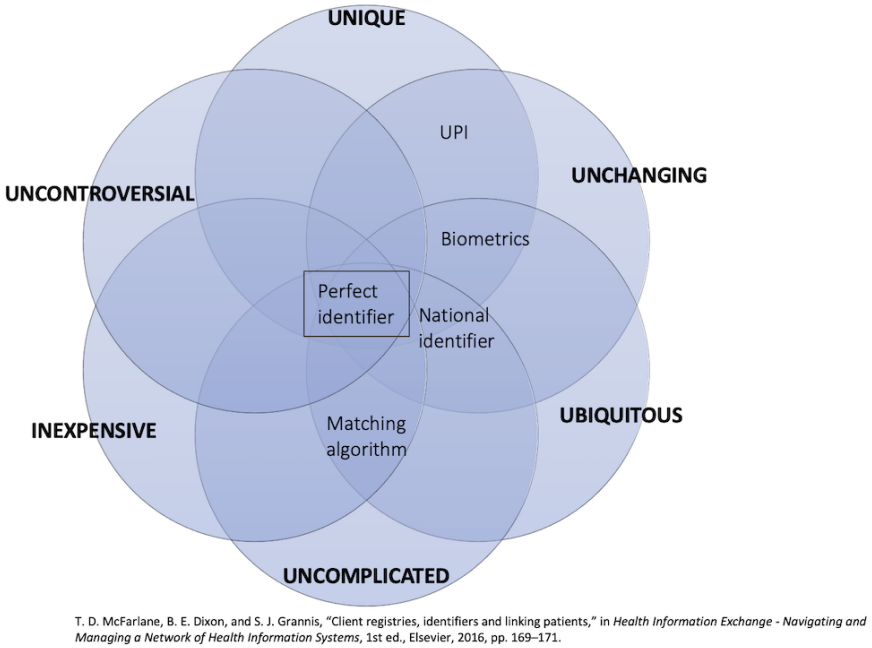


Figure 2.9: A framework for atributes and trade-offs of Patient IDs [265]

A national ID scheme is a dedicated repository of citizens' biometric information. Many countries have already established a national ID numbering system and database supported by the accompanying legislation. These systems are often designed to hold citizens' records and ensure they can be used across multiple sectors. Instead of creating a health-specific functional ID system, some countries have chosen to use the existing foundational ID systems, such as population registers or national ID systems, as the basis for patient identification. Estonia is one of the success stories for the use of national ID for health service and electronic prescription management [266]. Also, India is in the early stages of piloting the use of the Aadhaar national ID for health services [267]. National ID schemes often provide 'enrollment points' where citizens/patients can formally register and have their biometrics and other information captured into the national ID reposi-

tory. Most times, these enrollment points are different from healthcare institutions where health service is provided. Rather, a central state entity manages the enrollment, management, and issuance of a unique identifier. Depending on their design, these systems capture different biometric features like fingerprint, palm print, eye retina scan, voice, facial scan, and others to ensure uniqueness. These biometrics, in turn, aid cross-institution (or department) data matching, aggregation, and retrieval.

In some countries, especially in sub-Saharan Africa, patient ID is generated and managed by the institutions where patients receive healthcare [61]. Often, these institutional identification systems are not transferable outside the relevant institution, which could range from hospitals and primary health centers to insurance companies. These types of schemes are popular among paper-based health institutions. For countries with relatively little regulation around patient IDs, when a health institution transitions from a paper-based to a digitized system, it is common for software vendors to design and deliver institution-specific patient ID numbering schemes with the software.

A Master Patient Index (MPI), ensures that a patient within a health enterprise is represented only one time. It often leverage a dedicated software that assigns unique ID to each Patient within the enterprise, often across multiple health organizations [61]. Depending on their design, MPIs can have sub-patient indexes and can also incorporate national IDs as part of their dataset [62]. MPI is often centrally governed, and the central database facilitates duplicate management and issuance. It may appear trivial to discuss the question of who will be responsible for managing the MPI. However, in many countries (such as the US and Nigeria), state government plays a critical role in all areas that are not exclusive to the federal government.

Moreover, in most countries, in addition to the health ministries, there are a number of other agencies and regulators that have authority over different aspects of healthcare services. Matters are also further complicated by the mix of public and private stakeholders, coupled with different health institutions: from primary healthcare providers, through insurance service providers, to laboratory and pharmacy service providers. For such a system to be practical and functional, it must be designed to work effectively in most healthcare institutions. For instance, while a fingerprint-based MPI may be effective in uniquely identifying individuals, in practice, it can be difficult and expensive for each health institution to verify clients at service points [61]. Also, in my previous paper, the telecommunications providers' Subscriber Identification Module (SIM) numbers was proposed for use as a functional ID to facilitate Patient identification. Blockchain-facilitated

decentralized identity management has already been discussed.

2.4.2 Shared Health Records (Repositories)

HIE data storage depends on what architecture the participants choose. For a centralized architecture, the HIE provider will host the shared record of the patient. Patient-controlled models are designed for patients to either carry their data and share along or just control access to the data. Data can be on the patient's device or stored in a remote location. For Federated architecture, patient data is distributed in institutions that generate the data but provide access based on sharing agreement. In the decentralized architecture of HIE, data is accessible through the decentralized network even when the generating institution is not reachable. The scalability of a blockchain system or framework is largely dependent on the data storage strategy. Blockchain performance has yet to match the traditional central Application Programming Interface (API) models of say Visa card. For instance, data storage on the Ethereum blockchain is limited by gas price. Also, the bitcoin network transaction data cap limits its scalability. Most blockchain-based solutions like BlocHIE [128] use two storage strategies - on-chain and off-chain storage. On the Hyperledger fabric network, implementers can use either leveldb or couchdb, and there is technically no limit, except that imposed by signature aggregation and transaction submission time of the system.

2.4.3 Healthcare Revenue and Tokens

None of the healthcare blockchain-based papers discussed revenue models, even though the cost of implementing blockchain systems remains a major bottleneck to scalability. For instance, Xia et al. provided a detailed cost in space units for storing data on the blockchain network [138]. When 50 random participants used blockchain nodes on AWS over nine days, each participant was found to spend $ 283.85 per transaction to store all 259.2KB data on the Ethereum main net [232]. In addition, Li et al. also investigated the cost of data storage with a similar result [241] .

In addition, non-blockchain tokens have been implemented for health services in both developing and developed countries [67, 68]. In a blockchain network, tokens can be minted (generated), owned as a representation of a store of value, or transferred between different parties. Tokens are used for transaction fees on Ethereum and Bitcoin blockchain networks, and the Hyperledger fabric framework is extensible to support tokens. Decen-

tralized tokens have other uses like incentives, assets Security, transaction fees, utility, and governance.

2.5 Summary

This chapter discussed the systematic review identifying global HIE initiatives, including the state of HIE globally. The main barriers to health information sharing were technical, motivational, economic, political, legal, and ethical barriers. The evidence base of HIE, including the uses, effectiveness, and health outcomes of HIE for care collaboration and coordination, was presented. The literature evidence of models of HIE concerning data storage and service provision, including Centralized, Patient-centered, Federated, and Decentralized, were discussed. A summary of different identity management strategies was presented. This chapter also discussed the different organization's CA infrastructures: Traditional centralized-CA, one-trusted-CA, Multi-CA, or Client-CA. The survey also shows FHIR as the most used standard for data communication. Similarly, SNOMED CT was found to have the largest number of clinical concepts. No evidence suggests that SNOMED CT is the most used syntactic standard. Evidence of the cost implication of implementing a blockchain-based system and how it affects scale was also discussed. We have already published a systematic review of blockchain in healthcare and HIE, and it was well-received, with over 10,000 full-text views and 100 citations in the last two years. Also, my published book chapter on the role of digital ID in healthcare captured the traditional identity management strategies section.

The next chapter will discuss the Thesis methodology, starting with country surveys. The network design and reference implementation RegistryChain modeling will also be presented next.

3 Methodology

This section focuses on the methodologies used in this thesis, which include, Systematic literature search and analysis, country health facility mapping, country practitioner survey, and the development of RFD framework design.

3.1 Literature search & analysis strategy

A systematic search of existing literature was conducted to ascertain the state of the art in Health Information Exchange (HIE) research globally using the phrase "Health Information Exchange". The survey was conducted in three stages as follows: a) review of HIE globally, b) review of the state of the art of blockchain in healthcare, c) review of Hyperledger applications in healthcare. Each systematic search and review process is represented in Figure 2.1 to Figure 2.4 using the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) framework [268].

3.1.1 Systematic search of Health Information Exchange (HIE)

Separate searches and reviews were conducted for the global state of Health Information Exchange (HIE) over the last five years (2017 - 2021). Eight scholarly databases that capture the multi-disciplinary (Health and ICT) nature of digital health were searched. The databases were IEEEXplore, Ebsco, Scopus, PubMed, WebOfSc., ACM DL, SpringerLink, and Sci.Direct. The search helped ascertain the geographic spread of HIE adoption, the barriers to adoption, the standards used, and the architecture discussed in the reviewed papers. The phrase "Health Information Exchange" was queried in the databases, with the intention of filtering only technical articles. Initial queries returned about 18,000 items, and after title screening, only 1519 publications by title had relevance to HIE or health interoperability. Abstract review of all 1,215 articles reveals that only 210 had the quality or relevance for inclusion for full-text scheme-through. The 210 articles were subsequently Fulltext reviewed and formed part of the literature analysis in chapter two. The PRISMA

chart for this search and analysis is in Figure 3.1.

3.1.2 Systematic search of Blockchain in Healthcare and HIE

Early in the study, blockchain's potential to provide trustless intermediation between participants in healthcare was identified. A comprehensive literature search to determine the state of the art of blockchain in healthcare was conducted by searching eight scholarly databases: IEEEXplore, EBSCO, Scopus, PubMed, Web of Knowledge, ACM digital library, SpringerLink, and Science Direct. The EBSCO database provided access to 26 databases, including PsycINFO, MEDLINE, and Cochrane. The search conducted in the last week of May 2019 covered articles published between 2010 and 2019. This was the first search and publication done as part of this study. It has long been leant that most relevant data are in the last five years. The approach was to get comprehensive coverage of how blockchain is used in healthcare using keywords 'blockchain' AND 'health' or synonyms. The PRISMA for this search and review is in Figure 3.2. The findings were grouped into proposals, prototypes, and implementation. The few prototypes (n = 54) and implementations (n = 7) were full-text analyzed for performance, health standards and ontology used, security, privacy, cost, and storage schemes. The comprehensive literature search was published in February 2020 here [1].

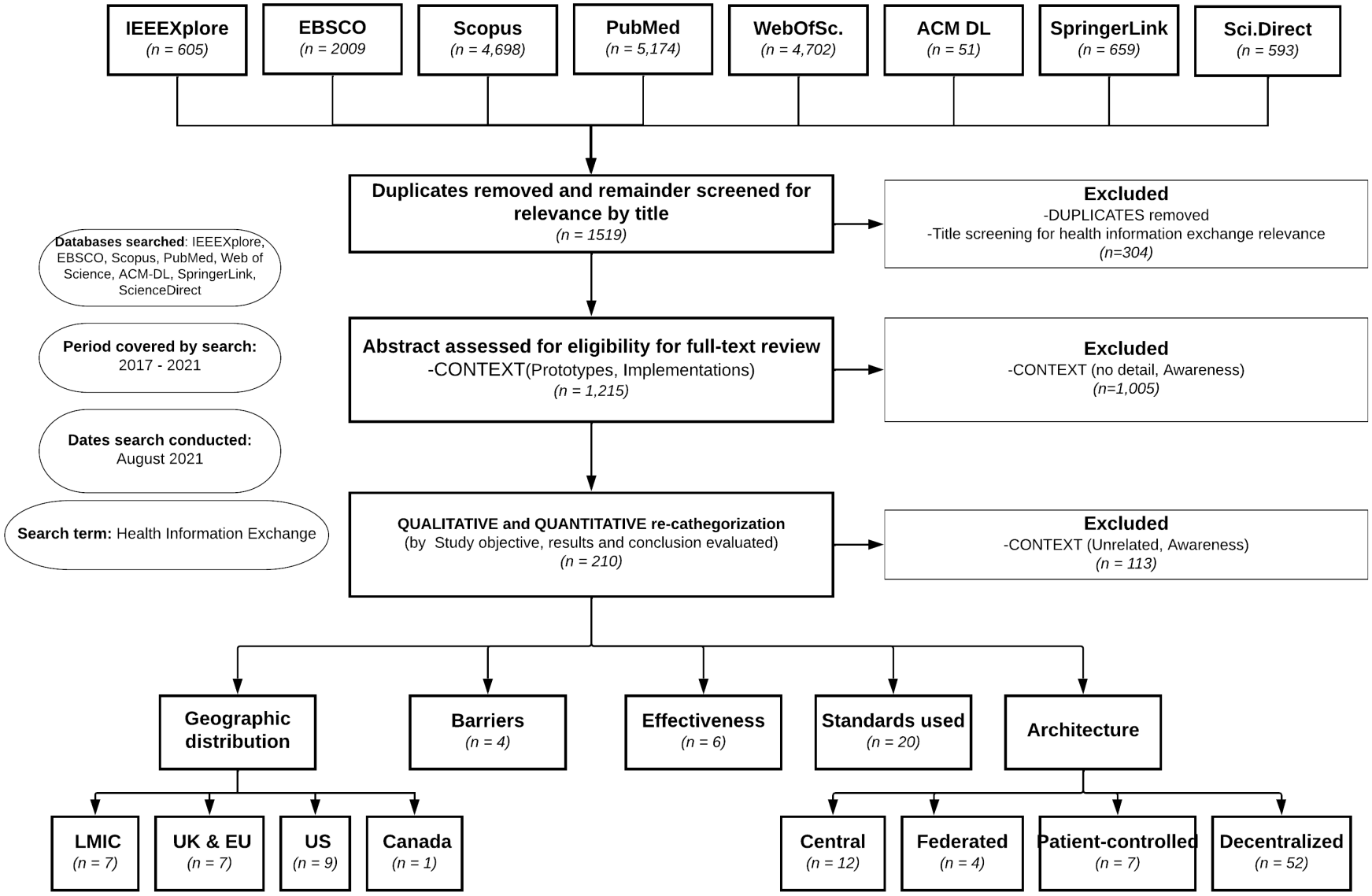


Figure 3.1: The PRISMA of review of HIE

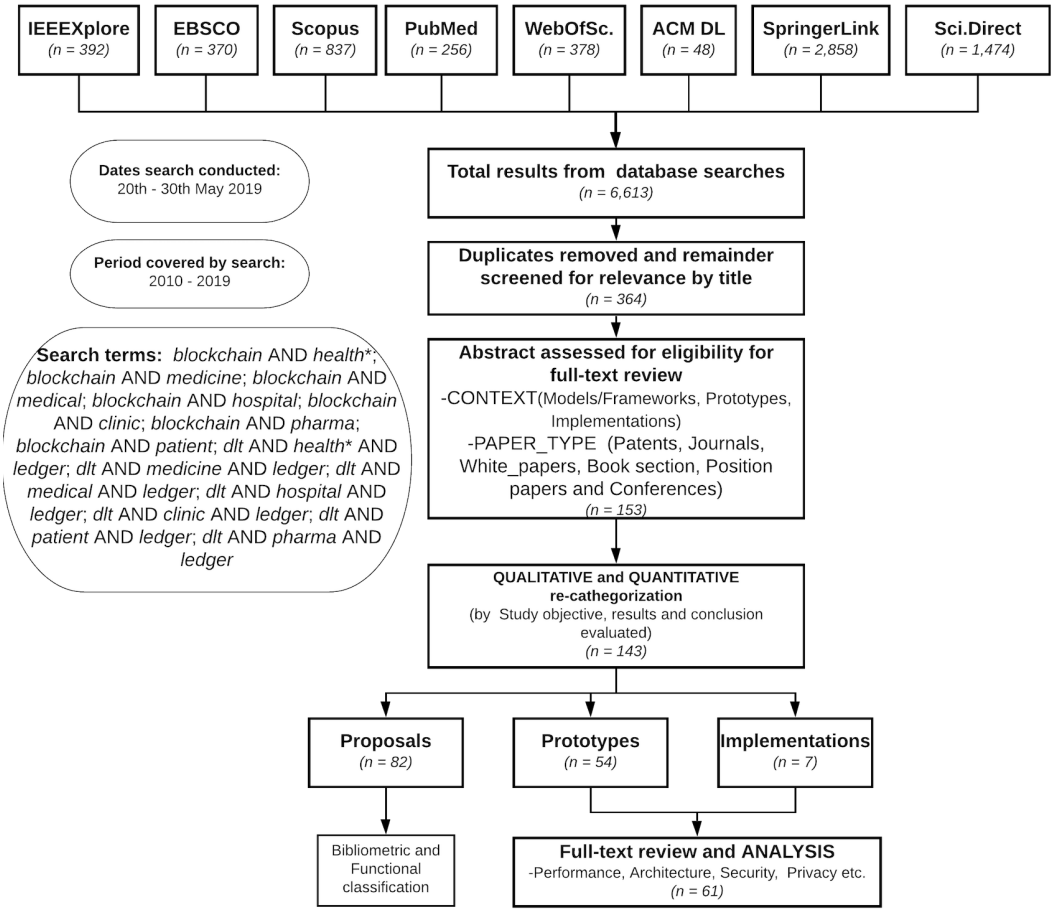


Figure 3.2: The PRISMA of systematic review of Blockchain in Healthcare [1]

3.1.3 Systematic search of Hyperledger in Healthcare and HIE

There are two broad categorizations: permissioned (eg. Hyperledger Fabric) and permissionless (e.g., Bitcoin and Ethereum) [1]. The permissionless blockchain also called the public blockchain, uses a wasteful Proof-of-Work (PoW) mechanism to reach consensus amongst parties. In August 2020, the cost for permissionless blockchain was equally modeled, simulated, and published for COVID 19 contact tracing [269]. The findings confirm the unsuitability of public blockchain for healthcare. Also, public blockchains require significant efforts to implement traditional regulatory functions. To my knowledge, no such regulatory function has been implemented in healthcare either by permissioned or by permissionless blockchain. The research was refocused on permissioned blockchain in healthcare because of better anonymity provided due to participant identity facilitated

security layer. Hyperledger is the leading enterprise open-source blockchain led by the Linux Foundation. In July 2021, a systematic search and review of the IEEEXplorer for articles published in the last five years (2017 - 2021) focused on Hyperledger in healthcare was conducted. Only the IEEE database was chosen as that is the only database with a relevant level of technical detail for Hyperledger-specific implementations. First, understand the literature scale around Hyperledger from technical journals and conferences. Second, to identify any gap in the literature between our 2019 survey and now related to blockchain in healthcare. Fifty-six (56) results were returned from searching IEEEXplore, and the abstract of all 56 publications was reviewed. Fifteen (15) articles were fully reviewed as relevant to this thesis.

3.2 Health facilities survey

The systematic search of literature was augmented by conducting an in-depth survey of health facilities in a typical LMIC country to ascertain the state of information sharing. While Sierra Leone may not be representative of LMICs, it indicates information flow and helps support the literature finding of no HIE in LMIC. This mapping was used to gather requirements for the health system case study and the ontology in chapter 4. This mapping also addressed the current state of information sharing in a typical LMIC. The survey targeted a representative sample of health facilities spread across all 13 health districts in the countries. Similarly, the survey also aimed to address the state of standards and terminologies used for information sharing. Sierra Leone, a country of 7 million population was chosen as a typical LMIC for the health facility mapping because of the researcher's digital health network in the country. We already published this survey [2].

A stratified sampling strategy was used, where all 14 healthcare districts were purposefully targeted for five or more health facilities. Seventy-two health facilities, consisting of 17 urban and 55 rural, were chosen from a total of 1,284 facilities across Sierra Leone. The margin of error is 11%, providing a 95% confidence level, for a representative sample. This was calculated using Calculator.net [270]. Ninety-six percent (n=69) of surveyed health facilities are public sector institutions. This very well aligns with the distribution outlined in Sierra Leone's National Digital Health Strategy and the state of health facility distribution previously detailed in our publication [2]. Health institutions were classified as urban/rural for spread and inclusion, based on information from the Ministry of Health and Sanitation (MoHS), in collaboration with the respective District Health Management Team (DHMT) heads. The health facilities were further classified according to the level

Table 3.1: Distribution of health facilities by districts [2]

District	No. in Rural	No. in Urban
Bo	4	2
Bombali	4	1
Bonthe	4	2
Kailahun	4	1
Kambia	5	0
Kenema	3	1
Koinadugu	5	0
Kono	4	1
Moyamba	5	1
Port loko	5	0
Pujehun	4	1
Tonkolili	4	1
Freetown (Wester Rural)	4	0
Freetown (Western Urban)	0	6

of their digital health activity into three: low, medium, or high digital health activity. Low, if there is no digital health solution; Medium when they have one or two digital health solutions; and High, when they have three or more digital health solutions. Given that each district has at least one higher level hospital, each district hospital was included in the survey irrespective of their digital health activity or wether they are in urban or rural location. This process was repeated until desired sample size of health facilities was reached for each district. Distribution of health facilities surveyed, by district is as in Table 3.1.

3.3 Health Practitioner survey

A survey of the current information-sharing process and datasets was equally commissioned to ascertain consistency of information sharing and help design the HIE model [3]. In healthcare, Referral-linkages are considered crucial for the proper functioning of health systems. A collaboration opportunity with a Federal University Ndufu in Ebonyi state Nigeria presented itself, and it was leveraged. Ebonyi state is a state in the Southeastern part of Nigeria with a projected population of 4 million. This survey was conducted to further establish the current state of standardization when information moves from multiple institutions. The referral use case was chosen as cross-institution referral is the most

Table 3.2: Distribution of respondents and their roles [3]

Typeofhealthinstitution	*Role*		
	Midewives	*Nurses*	*Doctors*
Primary Health Care (PHC) - clinics	3	3	4
General Hospitals - Secondary Healthcare	2	4	3
Teaching Hospital - Tertiary Healthcare	1	0	3
State Ministry of Health (SMoH)	1	0	0

popular use case for health information sharing.

Questionnaires were designed and sent to health professionals within the co-investigator at the university. Only 24 practitioners from 24 health institutions responded. The 24 practitioners from different health institutions all had different referral forms and different knowledge of what information is shared. The Fast Healthcare Interoperability Resource (FHIR) is an open standard. It facilitates structured storage and sharing of health information. Resources in FHIR are data structures representing an item in healthcare formatted in JSON or XML forms [3]. The developed FHIR profiles are these JSON or XML artifacts whose fields can be extended depending on each group's needs. Global best practice helps facilitate the standardized exchange of digital health information for better care. This survey used structured questionnaire to obtain information on information shared when a patient is referred-out or referred-in, the type of referral-forms used, the fields in the form, and the knowledge expectation for the client or their care-giver. The respondents were Physicians, Midwives, and Nurses, as in Table 3.2.

In addition, paper referral forms were reviewed, alongside the survey of the frontline health workers. The referral datasets from the forms were then mapped to profiled FHIR extensions. In the process, the data types and cardinalities, including resources and terminology binding were modelled. Questionnaires were sent out with a request for a copy of 'referral forms' to health workers in their respective health facilities in Ebonyi state, Nigeria. The survey was rolled-out between 10th and 17th June, 2019. Only 24 of the health workers completed and returned the questionnaires, while only three provided referral paper forms.

3.4 Network architecture design

The Regulated-Federated-Decentralized (RFD) framework was designed and presented with components and protocols. The three main components are the Distributed Applications (DApps) component, the external organizational components, and the Distributed Ledger Technology (DLT) component. The designed reference network architecture assumes a three-organization model run on docker containers nodes [271]. The choice of docker was to isolate components of the network and for its popularity. Also, docker is the primary isolation environment for distributing Hyperledger fabric blockchain, which reference implementation use [253]. As described earlier, the conceptual network architecture is composed of multiple organizations, and one such organization is shown in Figure 3.3. By definition, the organizational grouping boundaries are those with at least one CA. Also, it is possible to have multiple organizations using one CA; however, such uses will clearly be distinguished.

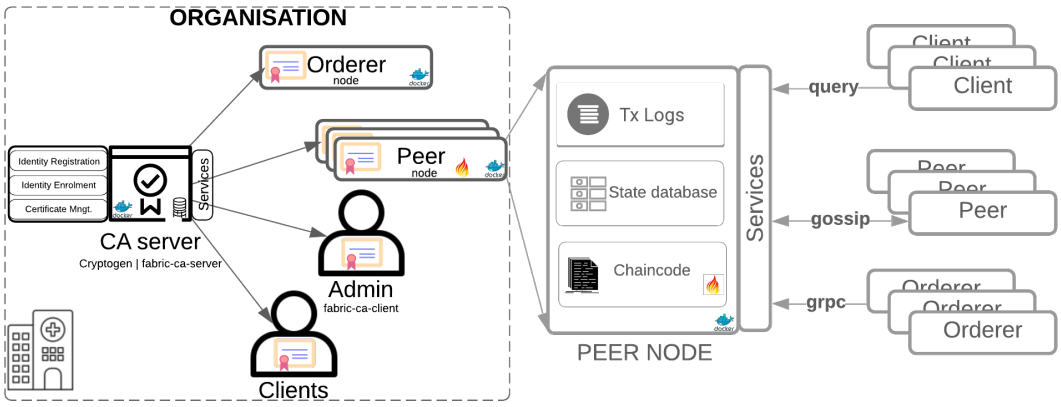


Figure 3.3: Organization with key components

The reference network configuration, code-named Reference RFD architecture for health registries change management (RegistryChain) is modeled using the FHIR ontology for DApps conformance and Hyperledger fabric blockchain [57], [253]. Each peer implements a FHIR R4 resource validation server in a Hyperledger fabric 2.0 peer node. Each organization has a CA and Orderer node for transaction block ordering, network submission, and committing. Tokens workflow was equally implemented in Smart contract.

3.5 Reference software design

User concerns were articulated in the *core.yaml* file used to configure each peer node. Also, business logic is represented using the UML and BPMN diagrams. The implementation used the Node.js stack, hence the FHIR Representational State Transfer protocol (REST) validator, the Hyperledger chain codes (smart contracts package), and client applications were configured for testing using Node.js.

3.5.0.1 FHIR REST API validator

The data structures for RegistryChain organization and individual assets was first modeled into four core FHIR structured bundles:

- IPS Repository
- PractitionerRegistry
- OrganizationRegistry
- TerminologyRegistry

Bundles are containers for FHIR resource [57]. The IPS FHIR resource was used for packaging and demonstrating exchanging a Patient's PHR [57]. The Practitioner FHIR resource is used to structure data from any category of *health workers' registry* on the RegistryChain blockchain. All categories of *organizations' registry* change can also be structured using the OrganizationRegistry FHIR bundle. Similarly, the *terminologies* are structured, changed, and managed using the same approach. The swagger specification of the reference server operation, modeled after core FHIR but minimally supported to demonstrate the select resource validation and FHIR conformance. For FHIR resource operation, the server does not implement message translation service or terminology translation. In order to keep it pluggable, it is assumed the organizations that traditionally undertake this service will agree on a structure before proceeding with data exchange. This model used the Node.js JSON schema validator.

3.5.0.2 Transactions

Smart contracts for each service may be implemented using Node.js Hyperledger-provided libraries. Implemented services are the 1) Organization attribute change management function, 2) Practitioner attribute change management function 3) Patient Identity manager, 4) Transaction endorser, 5) Transaction orderer, 6) Storage management, 7) Organization anchors, 8) CA provider and manager.

3.5.0.3 Security and privacy

The security of data exchanged using the reference architecture can be guaranteed in a number of ways: either the client enables a mutual Transport Layer Security (TLS) with the peer (server), or it can use a private data feature of Hyperledger fabric. The private data feature allows two or more parties to share information without other blockchain nodes (even endorsing nodes) seeing the actual data. The client can also optionally use PKI to encrypt the transaction data before submission to the blockchain peer.

3.5.0.4 RegistryChain Tokens Design

The business logic for the different tokens was implemented as smart contracts using Hyperledger Node.js Application programming Kit (APK).

3.6 The usage guide

The reference implementation was documented to generate a process guide for system end users. The process guide documented the hardware and software requirements and the steps for utilizing RegistryChain. The process guide provided where to access resources to extend RegistryChain.

3.7 Summary

This chapter detailed the overall approach country-level health facility survey, healthcare practitioner survey, and architecture and software design and evaluation strategies. Furthermore, the process guide for implementing the reference architecture was also presented. The next chapter will present the RFD framework and the reference healthcare implementation RegistryChain.

4 Results

This chapter presents survey findings and proposed model for health information to address the research questions.

4.1 Survey Findings

Here, the findings of the surveys conducted in support of the thesis is presented, one conducted in Sierra Leone, a mapping of health facilities. The findings from this work have already been published here by the researcher in collaboration with other researchers and government stakeholders from Sierra Leone's Ministry of Health [2]. Similarly, cross-institution information sharing survey of health practitioners in Ebonyi State Nigeria was conducted. We have published the details of the findings [3]. The findings of the health facility mapping and the health practitioner surveys will be presented in the subsections that follow.

4.1.1 Health Facility Mapping findings

In order to further understand the state of Health Information Exchange (HIE) by health facilities in a typical Low and Middle Income Countries (LMIC), a mapping was conducted in Sierra Leone with support from Sierra Leone's Ministry of Health [2]. Sierra Leone is a typical LMIC and provided the opportunity to validate the state of data sharing that exists at health facilities in an LMIC. Also, Sierra Leone is at the bottom of the global maternal health index in terms of outcome [272]. Given that most LMICs prioritized maternal health for improvement and digitization, Sierra Leone can be considered representative of a typical LMIC [272]. The mapping sampled 72 health facilities (out of 1284) in the country using a stratified sampling technique with a confidence level of 95% and 11% margin of error so that findings are generalizable. Sampling included at least five health facilities in each of the 13 health districts [273]. At least one hospital is selected in each of the 13 districts. The findings show that while many digital health solutions are spread across the

country, none shared individualized data critical for care coordination with other health facilities or institutions. For the hospitals, how they shared aggregated data is shown in Figure 4.1. The figure shows that 23% of hospitals share summaries (aggregates) with the district and the national health authorities. Another 23% share the summaries with the central health authorities only, and yet another 23% share with district authorities only. While 15% share the aggregates with the district, national health authorities, and NGO partners. Estimated eight percent share with no one, while another eight percent share with NGOs only. The majority only report (often paper curated) aggregate to the central government, while a few report aggregate information to the supporting Non Government Organization (NGO). See the details of this survey in our published work here [2].

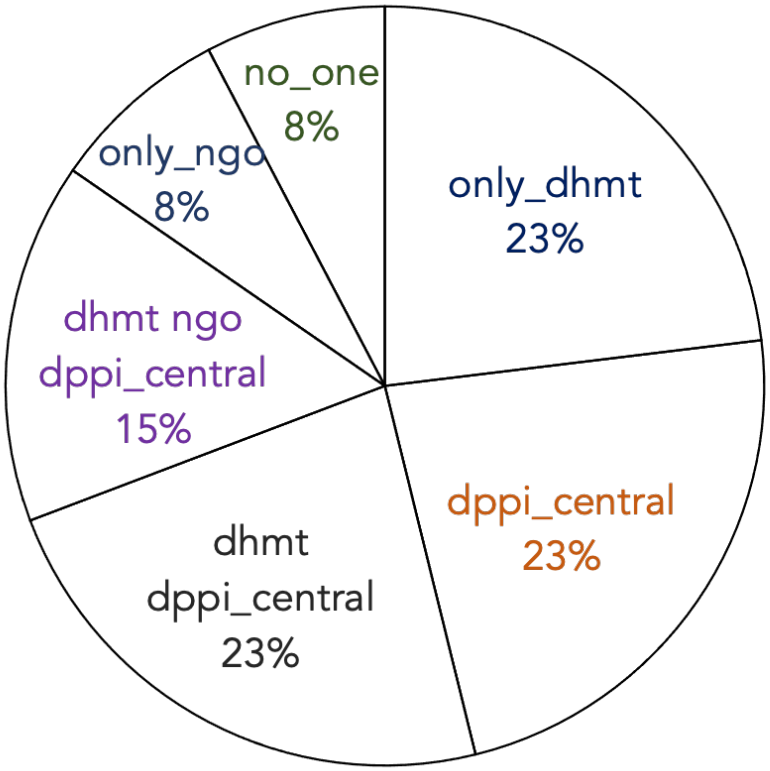


Figure 4.1: How hospitals share aggregate information in Sierra Leone [2]

4.1.2 Health Practitioner Survey finding

Responses from each practitioner varied significantly for all 24 of them [3]. For the three forms provided, they were mapped to survey questions to generate a single union of con-

tent of datasets. The responses regarding the datasets in the form varied widely and included other information not in the referral forms [3]. As part of this survey, already published, the Terminology and Individual resource groupings were prioritized for inclusion in the Fast Healthcare Interoperability Resource (FHIR) bundle using ICD-10 terminologies for maternal and child health chapters [3]. The image in Figure 4.2 show the datasets from the interview harmonization, profiled using the FHIR ontology. The primary FHIR resource used is the *Patient* resource. The image is from our earlier published paper, Chukwu et al. [3]. The data is presented in JSON format, one of the formats for Fast Healthcare Interoperability Resource (FHIR) structured health data [57].

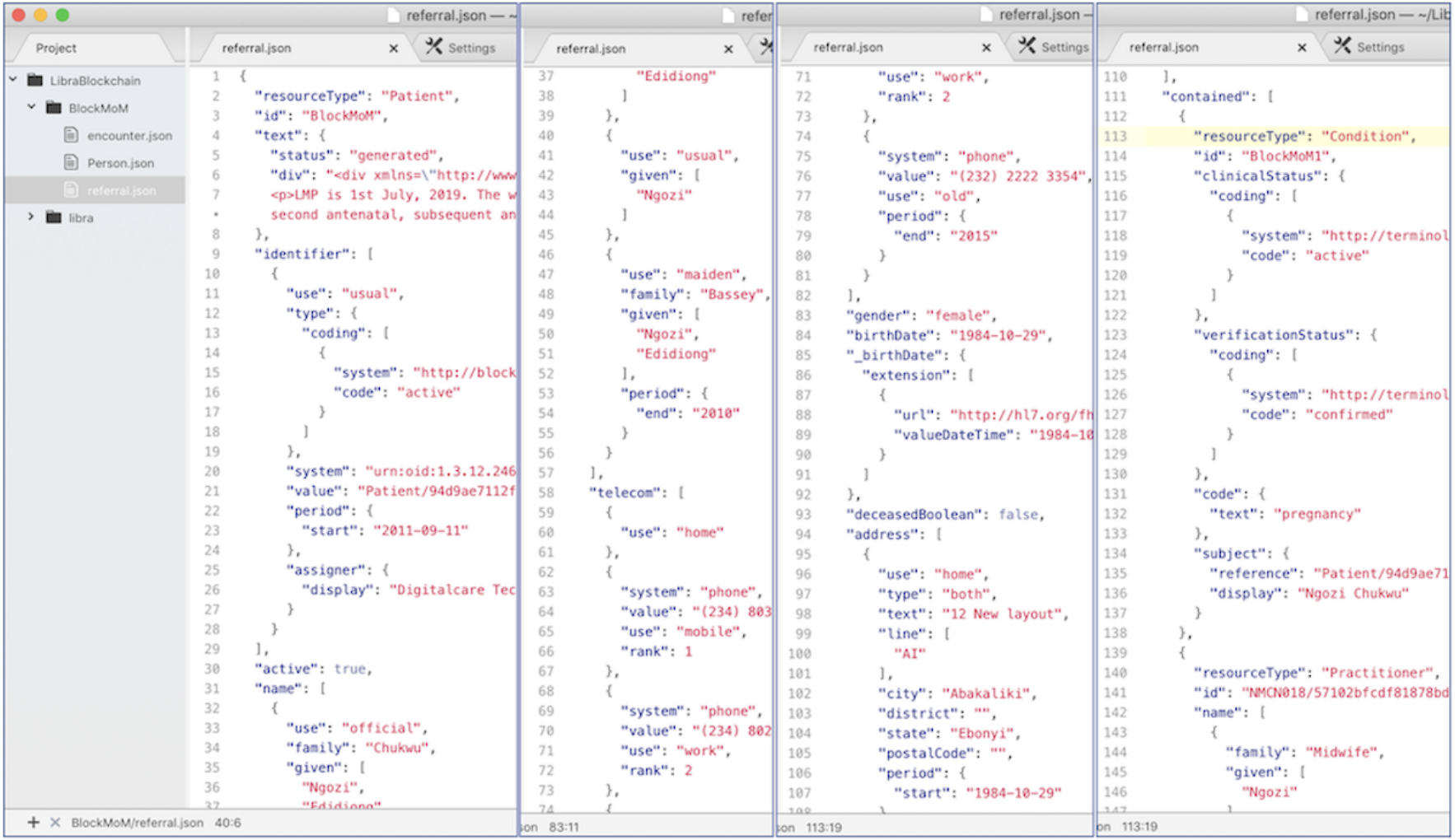


Figure 4.2: The profiled FHIR referral resource [3]

4.2 The Framework

The overall aim of this thesis is to design a framework for optimum information exchange. In this chapter, the novel framework, Regulated-Federated-Decentralized (RFD), for enterprise integration is presented. The RFD is an enterprise blockchain framework facilitated software integration model. The reference design and implementation of RFD named RegistryChain used for HIE is presented next. The logical and physical architectures, and the token algorithmic models were also presented. A high-level depiction of the RFD framework is illustrated in Figure 4.3.

Each organizational computing node A-F serves multiple clients. In Figure 4.3, simplistic organizational boundaries have been depicted using different color schemes to illustrate the RFD architecture. This hypothetical model shows three blockchain networks '1', '2', and '3'. Blockchain network '3' connects all the nodes. The first organization has nodes B, C, and F, the second, third, and fourth organizations have one node each depicted by different colors. The blockchain network 2 is joined by nodes E and F. The blockchain network 1 is joined by nodes A and F. In this framework, nodes A, E, and F serve clients within their organizations. So does node D that serve clients within its organization.

The users here can be either individual network participants or an application. The RFD model entails standardizing data ontology, terminology, and structures. It also includes a model for network tokens. As noted in Chapter one, the RFD framework will be illustrated using Health Information Exchange (HIE) use case called RegistryChain. The presentation of the overall framework developed as the main output of this thesis is discussed in each section: 1)The ontology and data structure, 2) RegistryChain, the reference HIE model, and 3)Token economics of RegistryChain. The user concerns were captured Participants, Functions (the basis for Transactions), Assets, and state attributes of the assets. The use case assets for shared services and data are shown in Figure 4.4.

Participants are Organizations, individual Patients, or Practitioners. The assets are also tied to a particular participant with an established role in the network. Each Participant's role is mapped to their function (or transaction). Assets are also characteristics of Tokens, Clinical concepts, Laboratory results, Medication characteristics, or Health summaries.

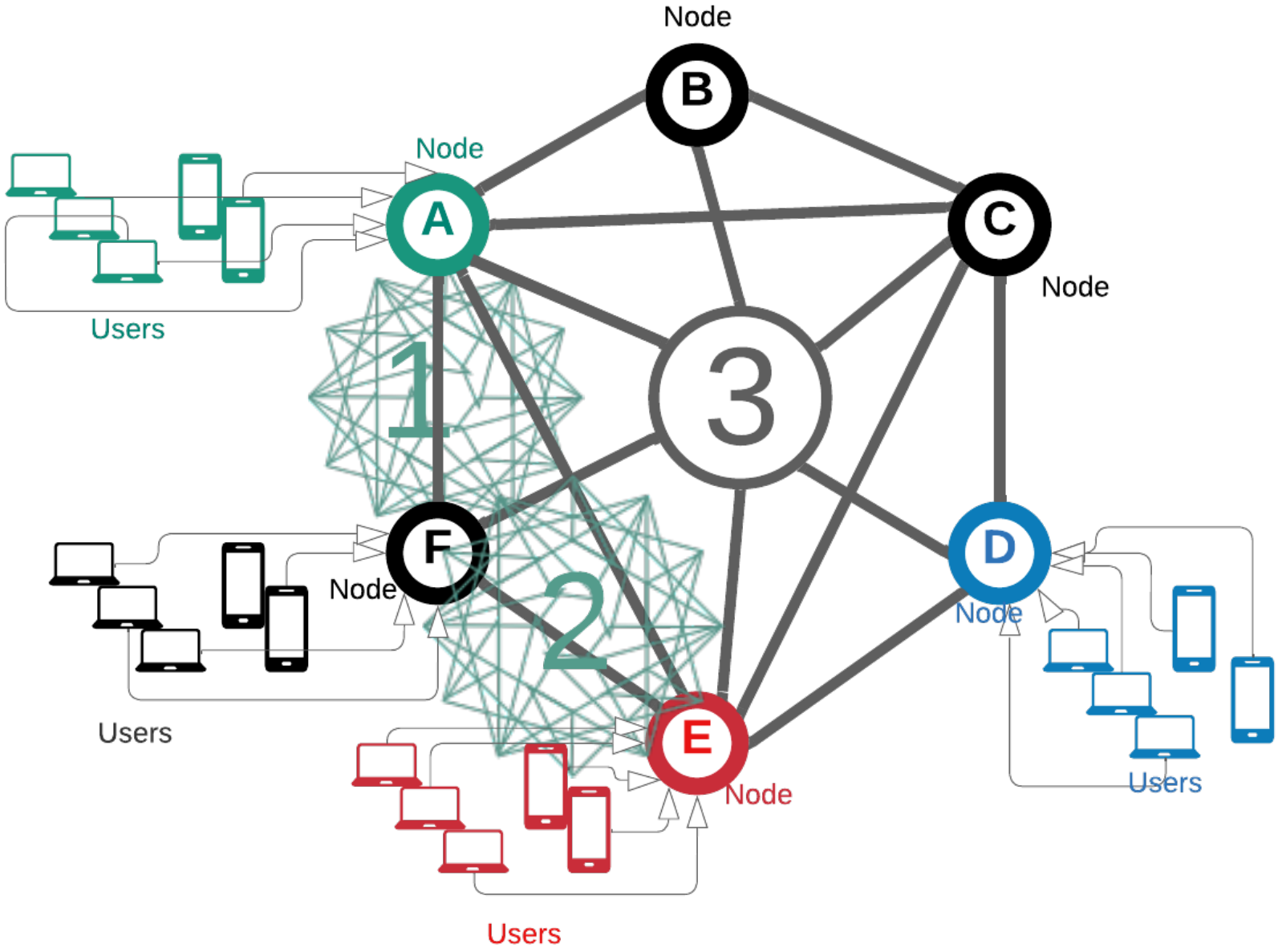


Figure 4.3: RFD shared services and repositories (self-drawn)

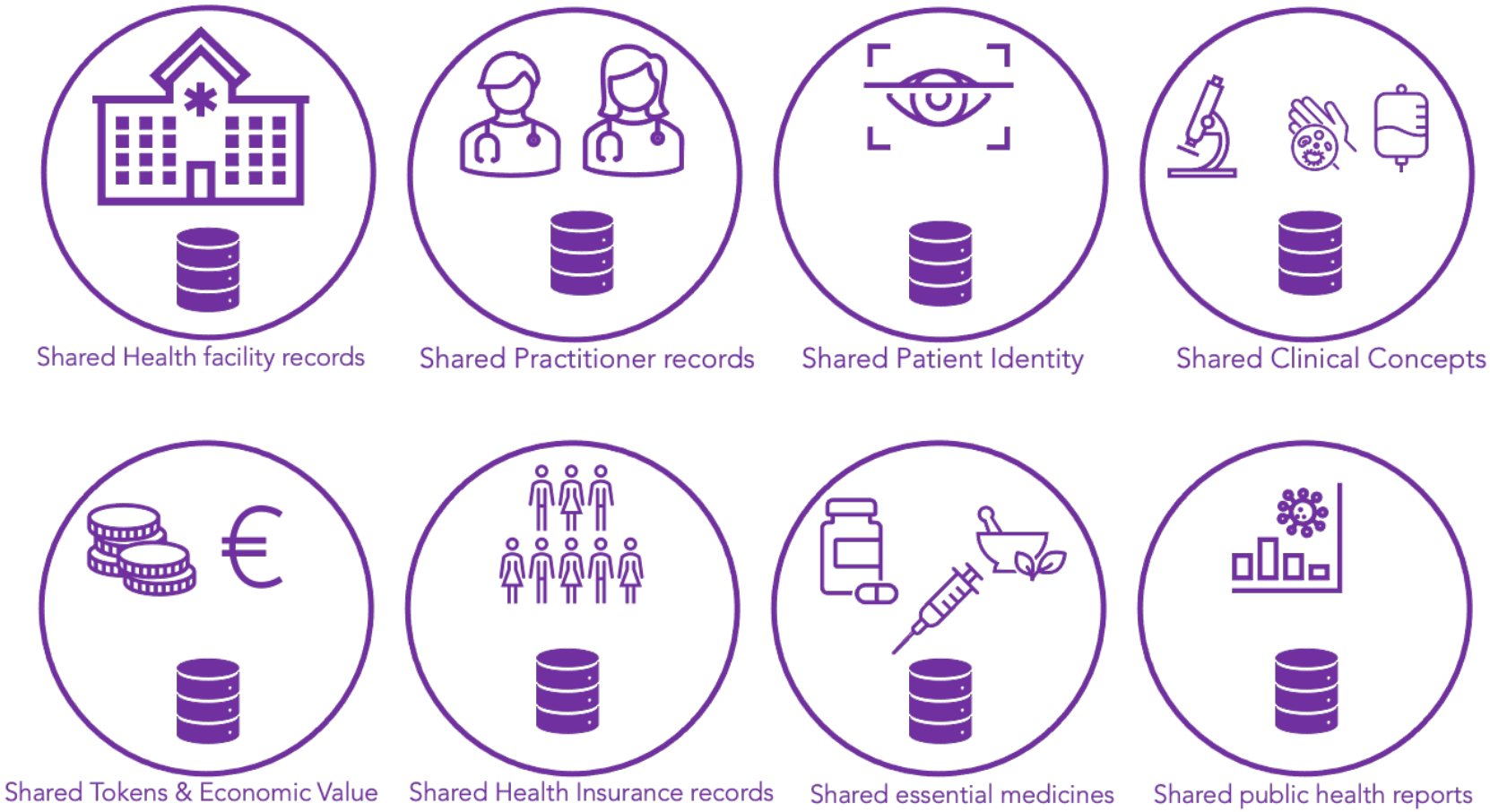


Figure 4.4: Example shared health services & repositories (self-drawn)

4.2.1 RFD Reference Ontology and Data structures

4.2.1.1 RegistryChain Ontology

RegistryChain architecture ontology represents the entities (or healthcare resources), descriptions, and their relationships. The base resources were modeled from FHIR resources and extended based on the thesis objectives [57]. The full RegistryChain Ontology for this use case is in Figure A.3. The CodeSystem ontology section is in Figure 4.5. The base FHIR ontology was adapted and extended as in Figure 4.5. The original FHIR resources in the ontology are highlighted in blue line, while the extended items are not.
The CodeSystem is a Fast Healthcare Interoperability Resource (FHIR) resource, and it is used to manage terminologies and registries. In this ontology, we extract the CodeSystem section of the designed ontology to show how other existing terminology registers fit in, notably SNOMED CT, ICD11, LOINC, RxNorm, ValueSet, and other less NamingSystems [12, 13, 14, 15, 16]. As noted in chapter 1, clinical terminologies can be classed broadly as poly-hierarchical and uni-hierarchical. The uni-hierarchical registers have terms (or concepts) that do not have a relationship with one another. These can also be called classifiers. The two main poly-hierarchical registries are SNOMED CT and RxNorm. The SNOMED CT is structured so that each concept has a class, a relationship class, and a description class. Similarly, the uni-hierarchical registries are made up of flat files, and ICD11 and LOINC are two main examples. Each row of data represents the concept information.

4.2.1.2 RegistryChain Data structure

Here, the attributes of the different bundles (group of FHIR resources) used for asset manipulation on RegistryChain are described. The attributes are structured using FHIR Revision 4.01. The data structure is made up of bundles which are arrays of FHIR resources represented in JSON format. The International Patient Summary (IPS) bundle as illustrated in Figure 4.6a include JSON formatted resources - $AllergyIntolerance$, $MedicationStatement$, $Condition$, and $ValueSet$. The PractitionerRegistry FHIR bundle in Figure 4.6b contain the following JSON formatted: $Practitioner$, $PractitionerRole$, $Organization$, $ValueSet$. Similarly, as in Figure 4.6c OrganizationRegistry bundle has $Practitioner$, $PractitionerRole$, $Organization$, and $ValueSet$. The TerminologyRegistry bundle as Figure 4.6d contain only CodeSystem FHIR resource, one for each of the code system used (e.g. one for SNOMED CT and one for ICD 10).

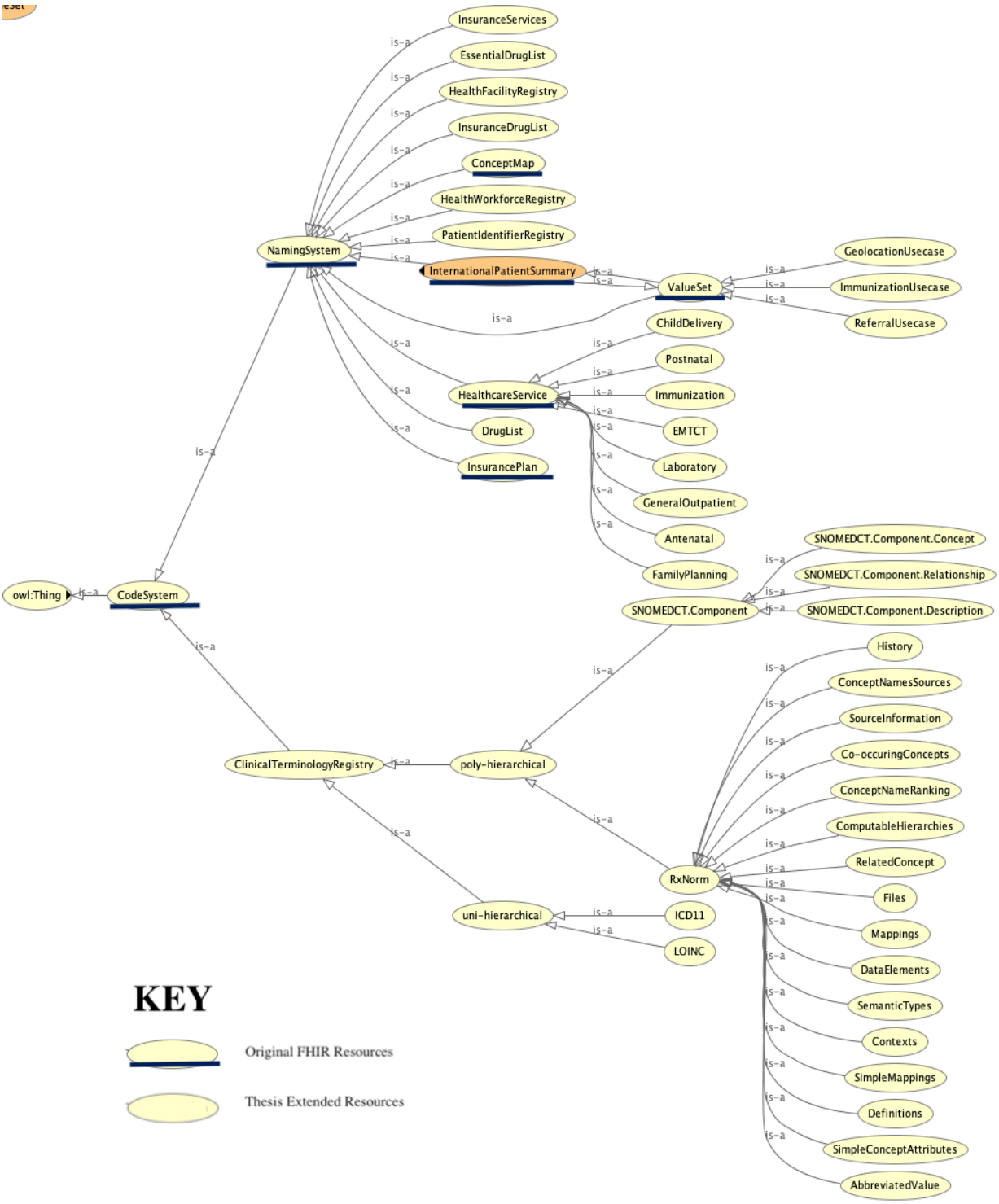


Figure 4.5: CodeSystem section of RegistryChain ontology (Prodégé generated)

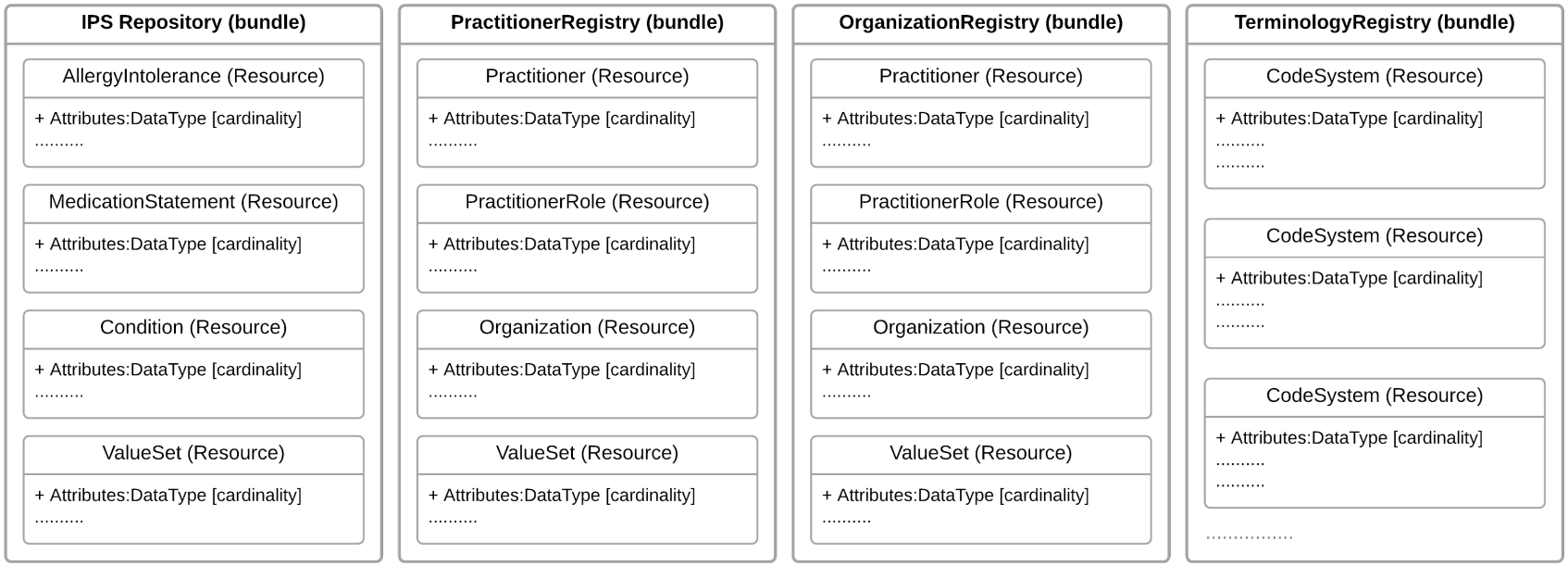


Figure 4.6: RegistryChain reference registries data-structures (self-drawn)

4.2.1.3 IPS Repository

The IPS is a bundle containing a number of FHIR resources grouped together for sharing Patient health information [57]. The IPS (v1.0.0 Standard for Trial Use (STU)1) based on FHIR R4 requires that the bundle, at a minimum, contain three resources as follows:

- Medication summary - Medication statement resource or Medication resource
- Allergies and Intolerances - Allergy Intolerance resource
- Problem list - Condition resource

The three mandatory resources in the International Patient Summary (IPS) bundle are the AllergyIntolerance, MedicationStatement, and Condition resources in Figure 4.6a. An IPS bundle can contain other recommended or optional resources as necessary. Immunization, Procedure, Organization, Performer, Observer, Device, Device Use Statement, Observation, Media observation, DiagnosticReport, Specimen, Imaging study, and Practitioner resources are recommended. Vital signs, Care plan, consent, and clinical impression FHIR resources are optional. In practice, the agreed content will be determined in advance by the jurisdiction and parties and the use case of interest. Only the three required IPS resources were used to generate the JSON bundle as objects for the blockchain world state for this simulation.

4.2.1.4 PractitionerRegistry

Similarly, the PatientRegistry, Organizational Registry, and Terminology Registry bundles are shown in Figure 4.6b to Figure 4.6d. The RegistryChain model provides for multiple Practitioners' Registry, which can be any of Pharmacy, Physician, Midwife, or any similar database managed by any of the other health and allied specialties in healthcare as shown in Figure 4.6b. A typical update to a Practitioner's detail includes role updates, the practitioner's detail updates, or their organization. Each of the possible attributes is covered using the combination of FHIR resources used in Figure 4.6b.

4.2.1.5 OrganizationRegistry

Similarly, organization regulators sometimes need to update the directory of staff in an organization, their role, or the details of services and other parameters in organizations (e.g., health facilities). The OrganizationRegistry bundle uses a combination of resources as in Figure 4.6c to capture these attributes.

4.2.1.6 TerminologyRegistry

FHIR represents terminology systems using the CodeSystem resource. The TerminologyRegistry bundle is used for packaging changes to CodeSystem or a combination thereof, as shown in Figure 4.6d.

4.2.2 RegistryChain HIE model

This section presents RegistryChain's logical and physical architectures, including identity management and the business logic coded into smart contracts.

4.2.2.1 Logical Architecture

The logical architecture is represented using The Open Group Architecture Framework (TOGAF) *Data Architecture* diagram. RegistryChain Network data architecture is based on a Hyperledger fabric permissioned blockchain. The novelty of RegistryChain is the facilitation and storage of shared health services (also known as registries) and shared health data (also known as repositories). The architecture supports data storage and uses tokens to facilitate shared value via a sharing economy. The high-level logical representation and respective functions of components of the RegistryChain architecture are shown in Figure 4.7. The three main components of RegistryChain reference architecture are the Distributed Applications (DApps) component, the external organizational components, and the Distributed Ledger Technology (DLT) component. The DApps provides a mechanism for accessing the blockchain system, while the external off-chain services and data storage represent the assets whose change is tracked on-chain. In the RegistryChain reference model of RFD, the patient repository (International Patient Summary (IPS)) is shown in Figure 4.7 on a blockchain world state. Similarly, OrganizationRegistry, TerminologyRegistry, and PractitionerRegistry attribute changes on the reference model were discussed. This model's smart contract encompasses the components' business logic as in Figure 4.7.

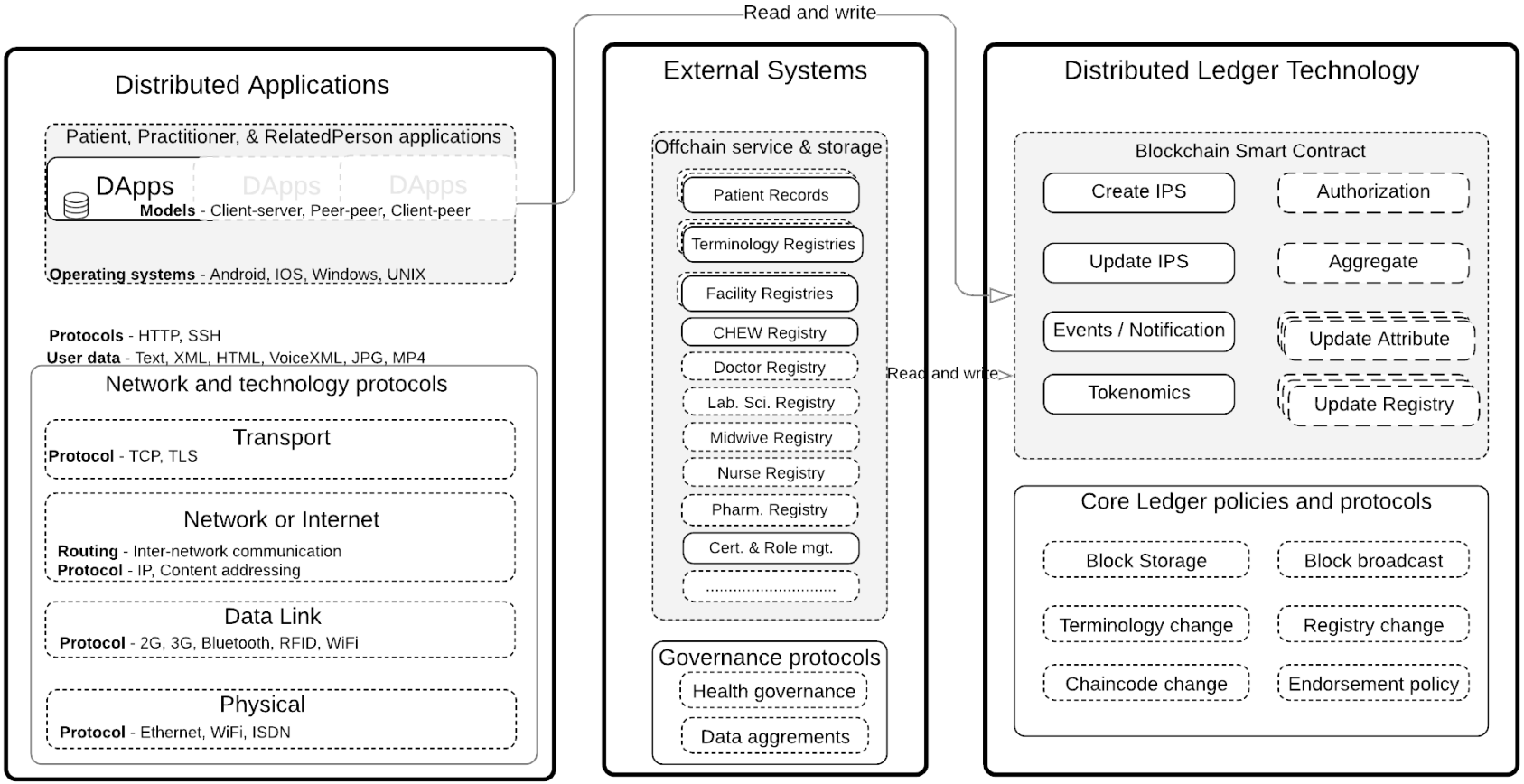


Figure 4.7: RegistryChain system components and protocols (self-drawn)

4.2.2.2 Physical architecture

The physical architecture is represented using The Open Group Architecture Framework (TOGAF) *"Technology Architecture"* components, one of the four TOGAF architecture diagram components. The TOGAF *"Technology Architecture"* is used to visually illustrate the technology components like network or software deployments. In order to better describe RegistryChain, its modularity, flexibility, and other features, we modeled a three organizations blockchain network. They are two software vendors V_1 and V_2, two healthcare regulators R_1 and R_2, and two health facilities F_1 and F_2. For simplicity, the RegistryChain model also sees *development organizations* as 'software vendors'. On the other hand, regulators are organizations with the regulatory mandate to create, modify, and approve respective registries. The blockchain network uses a multi-channel and multi-CA architecture. As in Figure 4.8, vendors V_1, V_2, and regulator R_1 are the CA providing organizations.

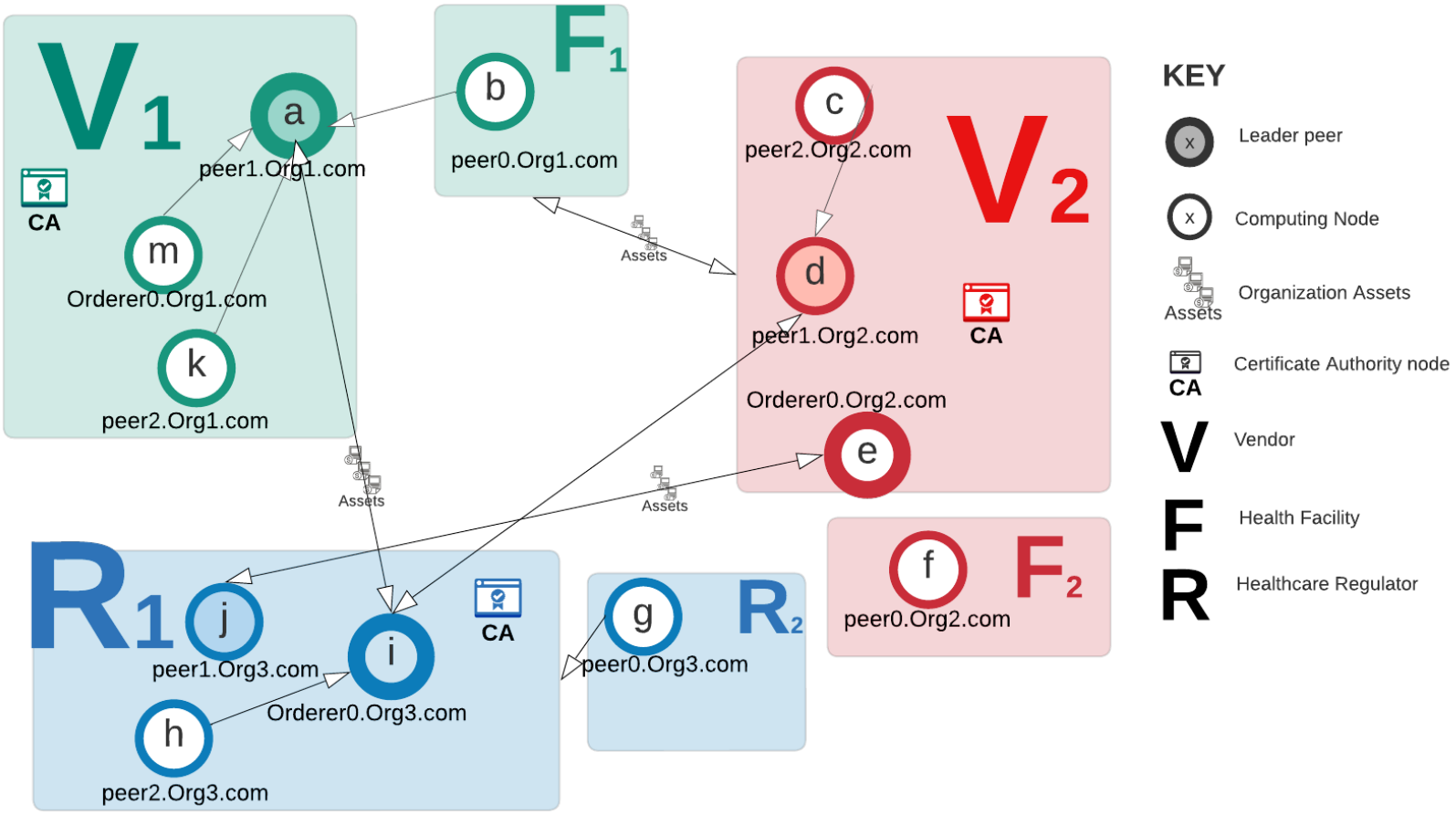


Figure 4.8: The RegistryChain Physical Network Architecture (self-drawn)

The CA right is depicted in the color scheme of red, green, and blue. Each color scheme also visually shows the identity scheme manager on the blockchain network. The CA nodes a, e, and i are responsible for the blockchain network's identity, key, and certificate generation, management, and revocation. The CA-providing organizations in this model has three nodes, each playing different roles as $Anchor$, $Orderer$, $Committing$, and CA nodes on the blockchain network. Notice that nodes g and m are used for the *Health worker registry change management* node. Changes approved by node g, for instance, to renew a Doctor's license, can be received by any node subscribed to that channel. Each asset to be committed to the world state is structured and validated for FHIR conformance for successful transactions. Though this model is used to illustrate RegistryChain in a hypothetical health system, it is by no means the only model supported by RegistryChain. Most of RegistryChain's architecture components are pluggable.

4.2.2.3 RegistryChain business process

The high-level business process of the reference model is presented here. The Open Group Architecture Framework (TOGAF) proposed *Business Architecture* as a high-level model for communicating business processes. Two widely used Universal Modeling Language Notation (UML) diagram for high level business process representation are *Use case* diagram and *Business Process Modeling Notation (BPMN)*. Also, the UML's Business Process Modeling Notation (BPMN) is a widely used artifact for high-level business process representation. It gained its appeal as a multi-stakeholder modeling notation that business executives and technical stakeholders can all relate to. The BPMN was used to represent the business process of RegistryChain as shown in Figure 4.9.

The BPMN diagram in Figure 4.9 illustrates three participants in each swimlane: Patient, Health Facility, and Regulator.

4.2.2.4 RegistryChain Identity model

Permissions in Hyperledger fabric are implemented using X.509 standardized Public Key Infrastructure (PKI) certificate. All identities within the network are issued certificates. These identities are 1) Individual participants (admins, users, and registrars), 2) Infrastructure components (Orderers, Peers, Apps), and 3) organizations (or members). Network participants, organizations, or infrastructure component can have more than one identity, and the identities are managed within their individual wallets. Each is issued a certificate containing information about the actor and their identity. The certificate has an attribute Rule for determining an actor's privilege level.Identity is generated in a two step process

1) invitation to enroll, and 2) enrollment. The registrar first initiates the process through an invitation to the end user to enroll by securely providing the user with an *Enrollment ID* and the *Enrollment secret*. Next, the user use these information to enroll and generates the certificates. The RegistryChain model leverages this framework for identity management on the network. The Hyperledger CA server has an embedded SQLite database for identity management, which may be replaced as needed.

The Hyperledger fabric CA has three main components - the CA server, the CA client, and the CA SDK. The server exposes REST API interfaces and services for certificate creation and management. The client makes available utility for interacting with the CA server REST API services. The SDK on the other hand make available libraries for writing applications that interact with the CA server. The admin and registrars can register, create and manage identities, manage affiliation, and revocation. Similarly, the end user can enroll and re-enroll as needed and permitted. The CA server expose the services as REST API, then consumed by either the CA client or the SDK in an application. The REST specification are available in swagger formats.

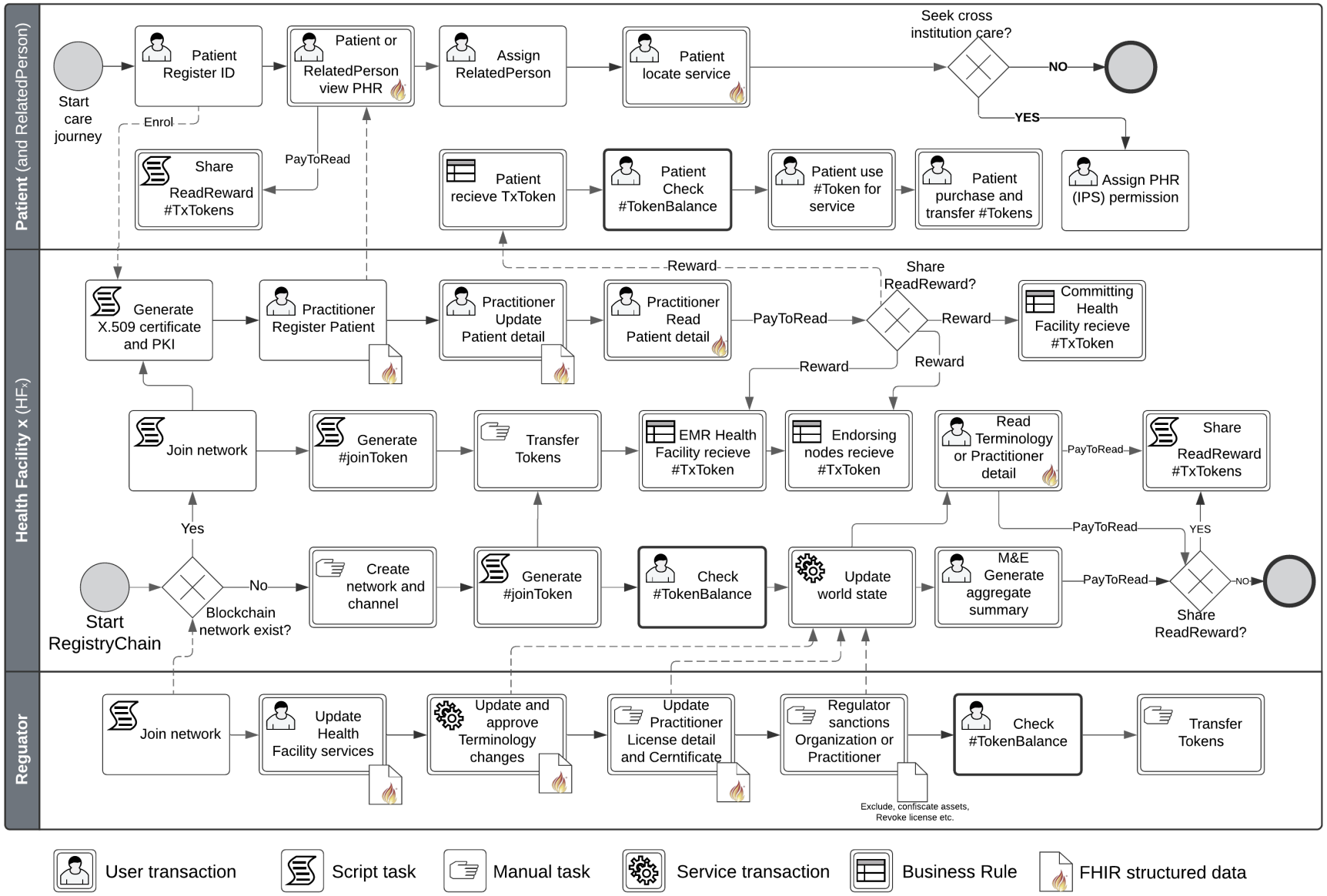


Figure 4.9: RegistryChain Business Process Modeling Notation (BPMN) (self-drawn)

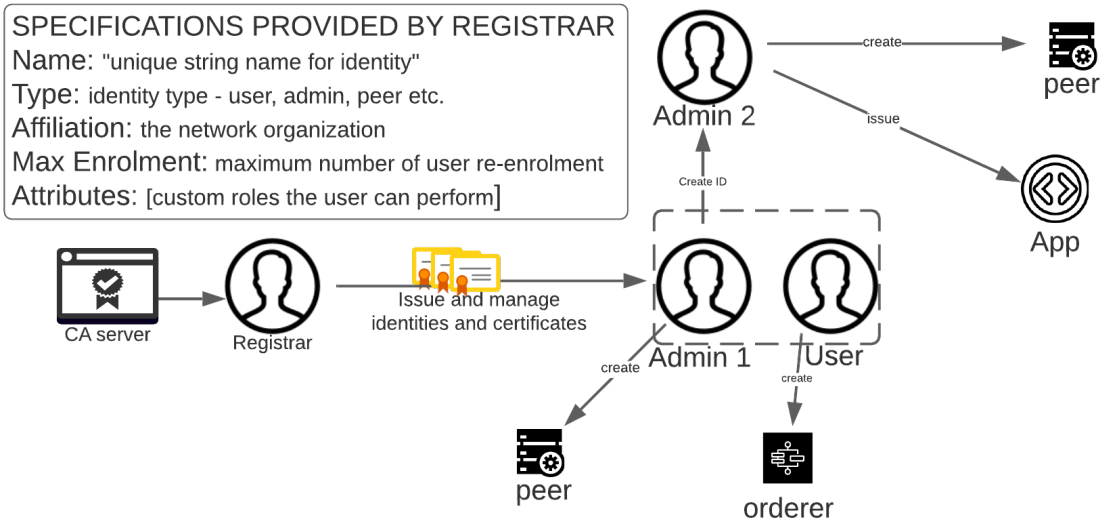


Figure 4.10: RegistryChain Identity management inspired by Hyperledger framework (self-drawn)

The Figure 4.10 highlights the hypothetical identity management for application of RegistryChain, this image was inspired by the Hyperledger fabric framework identity management workflow. Health facilities can create or join a network as they see fit as long as they have the docker container with RegistryChain binaries. In this model, the health facility hosts a mini Certificate Authority (CA) node responsible for generating and managing the network certificate and identity for self and Patients. Health facilities will issue themselves member-level identities and certificates and the patients' user-level identities and certificates. Similarly, Practitioners working in the health facility will be issued the facility-specific certificate and identity. In addition, health facilities are also members of an Organization Unit (OU) grouping as specified by the jurisdiction, say a state or a district. This means that each health facility will have at least two identities, one locally generated and the second generated by a higher OU CA.

Additionally, the Patient can enroll and obtain X.509 certificate and cryptographic PKI for any OU group network identity and asset signing on the blockchain. The Patient can view their FHIR structured PHR (and for the use case, the PHR is the IPS bundle) associated with any identity in their wallet. The Patient can assign permission to a RelatedPerson (a FHIR resource) to read and administer permissions to their PHR on their behalf. The Patient cannot update or change details of their PHR; Only an authorized Practitioner can

update Patient PHR as a quality control measure.

Authorized Practitioners can create a patient PHR by calling the appropriate smart contract functions to generate the Patient's identity using the health facility mini CA. The generated identity certificate is stored in the user sub-directory specified for the health facility. The Patient generated identity is not accessible to the Practitioner. A health facility-specific unique cryptographically generated and human-readable hash number can be used for search and Patient resource location. Additionally, the identification may use the mobile number attribute or a national identification attribute for local search and resource location update operations of the PHR. Such changes are logged on the blockchain network channel the organization (organization) is part of. Any network participant reading the Patient's PHR will pay in network tokens, which will then be distributed to eligible network members. The details are described in the *Registry Token Economics* section.

The regulator is on the network to oversee and ensure rules (as implemented in the smart contract and endorsement policies) compliance. Different categories of regulators can individually hold the right to update different asset types, as noted already- the TerminologyRegistry, the PractitionerRegistry, and OrganizationalRegistry are a few examples per the use case. As a network participant, the regulator can equally own tokens. However, the reference implementation, as in the BPMN diagram, does not include regulator minting tokens when they join the network.

4.2.3 RegistryChain Token Economics

Objective 4 is to design a token-based model for transactions and shared value economy on the blockchain network. How tokens are used on RegistryChain is described here. Like most DLT-facilitated systems, RegistryChain uses tokens to represent key assets. Tokens are used to provide a base value and unit of exchange for fungible assets. Assets that can be tokenized are Individual reputation points; Server uptime; Financial assets; License and certificate standing of provider; Services provided by health facilities, and combinations of these assets. Ethereum cost-to-token model is already established and tested for several machine transaction types and data sizes as detailed in Table 4.1 [4]. The Ethereum cost is measured in Gwei, which is a billionth of Ether (1Gwei = 0.0000000001ETH). RegistryChain will use the relative transaction gas cost to estimate eventual costs in practice for different transaction types. The model uses these relative rates as the benchmark as they illustrate the relative computational and storage costs.

Table 4.1: Transaction gas price estimation based out Ethereum pricing [4]

Transaction Type	Gas Price (Gwei)	Relative Price
Create new contract	32,000	16,000
Save a word to storage	5,000 - 20,000	2,500-10,000
Load a word from storage	200	100
Sign a transaction	1,000	500
Get an account balance	400	200
SHA3	30	15
Save a word to memory	3	1.5
Load a word from memory	3	1.5
Get an account address	2	1
EXP (exponential)	10	5
MUL (Multiplication)	5	2.5
DIV (Division)	5	2.5
ADD (Addition)	3	1.5
SUB (Subtraction)	3	1.5

The ERC20 (Ethereum Request for Comments 20) proposed by Fabian Vogelsteller in 2015 is the standard for creating tokens on the Ethereum blockchain platform in the form of smart contracts [274]. There are six mandatory and three optional functions that an ERC20 implementation needs to have to be deemed compliant [274].

The benefit of following a recognized tokenization standard is that tokens created with the ERC20 standard can be easily listed on existing exchanges. Wallets used by existing Distributed Applications (DApps) can be repurposed for any new blockchain and token with minimal modifications. Coding and maintaining tokens and the exposed interface will be easier. RegistryChain implements four of the nine ERC20 functions: $name$, $totalSupply$, $balanceOf$, and $transferFrom$. Other functions like $decimal$, $symbol$, $approve$, $allowance$, and $transfer$ [274]. For consistency, RegistryChain uses nomenclature already established for blockchain token administration. For instance, the token is denominated in the gas unit as used in Ethereum. See simplified Beige paper explanation here [4]. Though the gas token unit in RegistryChain is different from the gas token unit in Ethereum, it facilitates clarity. Next, tokens generation, use for transaction payment, and transaction transfers is discussed.

The reference implementation, RegistryChain proposes four transactions for tracking: 1) **Organisation-join** token mint, 2) **Integrity token minting award**, 3) **IPS asset commit (store)** & **IPS asset read**, and 4) **Token asset transfer**.

4.2.3.1 Organization-join Tokens Minting

The model assumes that when organizations join a given blockchain network using RegistryChain, each organization joining the consortium with approved Organisation Membership Service Provider (MSP) will mint a fixed number of cryptographic tokens O_t usable on the network. The first organization to bootstrap the network will mint the genesis token (first organization token), G_{nt}. At inception, when it is only one organization, G_{nt} = O_t. See Algorithm 1. When the next organization is approved to join as an organization-level MSP with ordering service capability, it will also mint O_t number of tokens based on $equation1$. At any given time, the total number of tokens minted due to organization-level joins, TJ_t can be computed using $equation2$. Also, before joining, an organization can compute their proposed equity (or stake) at join using $equation3$. Subsequent organizations joining with one MSP (and one endorsement node) will use these formulas. The actual number of tokens minted as a result of *organization-join* operation is fixed but varies by network. The recommendation is not to exceed the token equivalent of *One Gigabyte of data storage* estimated at $125thousandwords\ tokens$ (using $8bits$ size per word benchmark).

Algorithm 1: Organization-join ERC20 token minting algorithm

Input: MSPID, id={id_1,id_2, $\ldots id_n$ } ; $NumberOfGenesisTokens$, G_{nt} ;
$NumberOfExistingOrg$, EO_n; $TotalJoinTokens$, TJ_t
Output: $EquityAtJoin$, E_{aj}; $TokensPerOrg$, O_t; New $TotalJoinTokens$, TJ_t;
$SuccessStatus$, S_s
async **checkPermissionToMint**(id, G_{nt}) {
"Check for permission to mint"
return (AccessToMint)}
async **checkEquityAtJoin**(TJ_t, EO_n, G_{nt}) {
return(O_t, TJ_t, E_{aj})}
async **MintJoinToken** (id) {
"Mint tokens for new organization"
return (S_s, TJ_t)}
Update mint state
Emit events}

The detailed equation for computing $TokensPerOrg$, O_t, $TotalJoinTokens$, TJ_t and $EquityAtJoin$, E_{aj} is as in equations 1 to 4.

$$O_t = \frac{TJ_t}{EO_n} \tag{4.1}$$

$$TJ_t = O_t * EO_n \tag{4.2}$$

$$TJ_t = \sum_{i=0}^{EO_n} O_t \tag{4.3}$$

$$E_{aj} = \frac{1}{EO_n + 1} \tag{4.4}$$

The TOGAF *Application Architecture* component was used to visually show a three-view process for organization join in Figure 4.11.

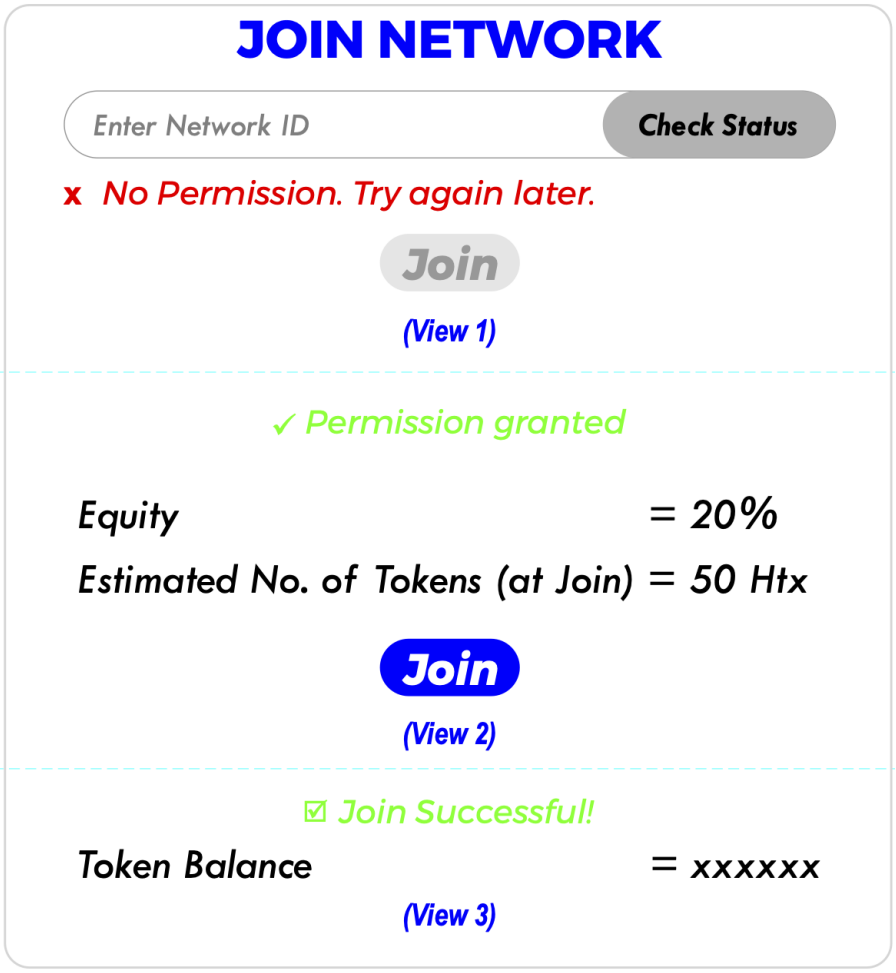


Figure 4.11: Organization-join, token mint workflow (self-drawn)

4.2.3.2 Integrity Tokens Minting

Algorithm 2 illustrates the Integrity threshold IT token mining process.

Algorithm 2: Integrity score ERC token minting algorithm

Input: KA_n ={KAS_{n1}ANDKAA_{n1},KAS_{n2}ANDKAA_{n2}, ... KAS_{nz}ANDKAA_{nz}}, $IntegrityTreshold$, IT

Output: $IntegrityScoreToken$, IS_t

async **computeIntegrityScore**(id, G_{nt},IT) {

IS_t = KA_n * IT return (IS_t)

This transaction happens when say, a committing node of an organization meets a given keep-alive, KA_n response threshold (e.g., 98%) computed from

$NumberOfDailyKeepAliveSent$, KAS_n, and

$NumberOfDailyKeepAliveAck$, KAA_n. Integrity threshold IT is best to set at smart contract deployment with agreed consortium members.

4.2.3.3 IPS FHIR bundle asset transfer

The International Patient Summary (IPS) asset is created and committed on a blockchain ledger through a series of steps as illustrated in the UML sequence diagram in Figure A.6.

The first step is performed off-chain by either the client application or the committing node to validate that incoming data for the commit proposal is in conformance to the set FHIR schema structure. In this case, this validation is performed by the source peer F_s using the *asymetrik* FHIR JSON schema validator npm package. Once the validation is passed, the next sequence will be to send a transaction request with the IPS bundle (containing JSON array) of FHIR resources to the endorsing nodes on the network channel. The nodes on the channel in the reference case in the sequence diagram are V_1, V_2, R_1, and R_2. F_s will gather all transaction payloads with endorsements from the majority endorsing nodes and proceed to post endorsement commit requests to authorized ordering nodes within their organization. The ordering node will check the transactions for requisite signatures, then order the transaction along with other transactions from the same organization up to a set limited quoter. The ordered transaction will be sent to other orderers across organization boundaries for commit endorsement. When the orderer receives all the commit signatures, the block will be committed, and replication will be initiated to other committing nodes using the gossip protocol. Commit and successful replication acknowledgment is sent back to the source facility F_s from all committing nodes that receive commits, including F_D. Updating an IPS record will follow this same approach.

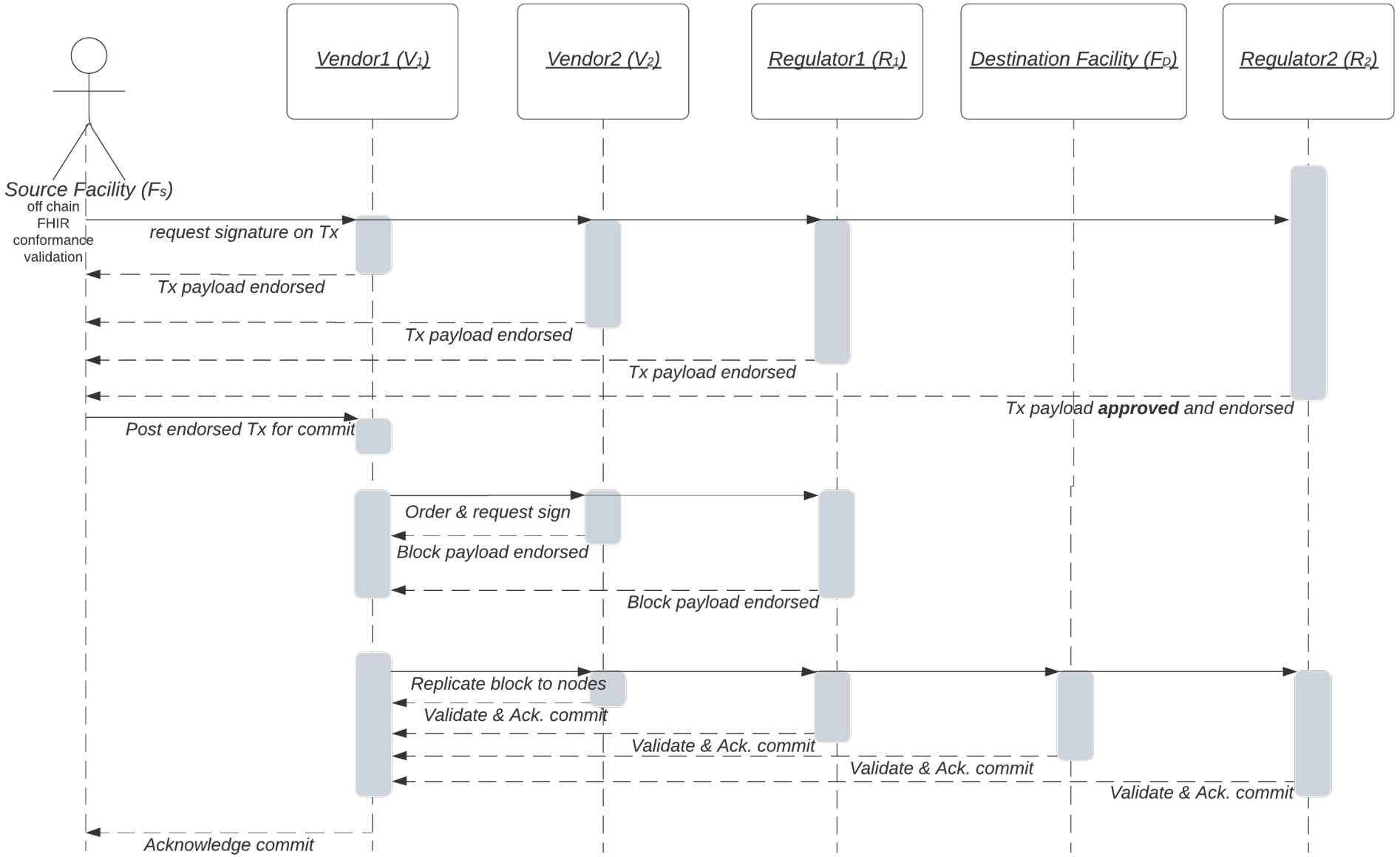


Figure 4.12: The blockchain commit UML sequence diagram (self-drawn)

A typical example is when a researcher requests an aggregate of world state data. The requested transaction will be paid for in a blockchain token. The algorithm for this is in Algorithm 3.

Algorithm 3: IPS bundle asset transfer smart contract algorithm

Input: $SourceID$, id_s; $PatientID$, id_p; $EndorserIDs$, $id_e = \{id_{e1}, id_{e2}, \ldots, id_{en}\}$; $CommitterIDs$, $id_c = \{id_{c1}, id_{c2}, \ldots, id_{cn}\}$; $CommitEnvelopDataSize$, D_{size}, $DataQualityIndex$, D_{qi}, $ReadDataFee$, R_{fee}, $TypeOfTransaction$, T_T

Output: $CommitterIDsWithPatientData$, $id_{cp} = \{id_{cp1}, id_{cp2}, \ldots, id_{cpn}\}$; $DataSize$, D_{size}, $SuccessArray$, S_a, $NumberOfEndorsingNodes$, EN_n

async **commitTxToken** (id_s, E_{aj}, id_p, id_e, id_c) {
"Check endorsement policy (outside smart contract)"
"Check IDs (outside smart contract)"
"Check data structure & FHIR conformance (outside smart contract)"
"Check transaction type (eg. FunctionCall, MemoryUse, and DataSize)"
"Execute Commit transaction"
return (id_{cp}, D_{size}, EN_n)}

async **ReadTxToken**(id_s, E_{aj}, id_p, id_e, id_c, R_{fee}) {
"Check endorsement policy (performed outside smart contract)"
"Check IDs (performed outside smart contract)"
"Check T_T (eg. Computation, FunctionCall, MemoryUse)"
"Check executor token balance"
"Check associated fee"
"Compare balance vs fee"
"Charge fee"
"Assign pre-determined tx fee sharing to participants id_s, E_{aj}, id_p, id_e, id_c"
"Execute Read transaction"
return (S_a, EN_n)}

Tokens may be exchanged when data read transactions happen on RegistryChain. The relative amount of tokens exchanged for a read transaction of committed data will vary and depend on node participation, the type of the transaction, the size of committed data, and if a refund was paid when a qualifying storage freeing transaction is executed. When an endorsing node participates in a transaction (meaning they signed a transaction), they earn an equivalent token, paid for by the transaction beneficiary. The thesis already shows from Table 4.1 the role *transaction type* plays in the model. The thesis assumes that the majority of token fees will fall under one or more of the following RegistryChain transactions.

- Computation fee

- Function call fee
- Increased memory usage fee
- size of data requested fee

4.2.4 Tokens: transfers and fiat conversion

Like most cryptocurrency tokens, tokens implemented on RegistryChain are transferable and usable on the network. In addition to payment for transactions, token holders can sell or share their tokens with individuals or organizations with valid CA-issued identity on the blockchain.

Algorithm 4: Token transfer smart contract algorithm

Input: $SourceID$, id_s; $RecipientID$, id_r; $TokenBalance$, T_b, $TokenTransferValue$, T_{tv}, $EndorserIDs$, $id_e = \{id_{e1}, id_{e2}, \ldots, id_{en}\}$; $TypeOfTransaction$, T_T
Output: $TokenBalance$, T_b; $SuccessArray$, S_a
async **TransferToken**(id_s, id_r, id_e, T_T, Tx_{fee}) {
    "Check endorsement policy (performed outside smart contract)"
    "Check T_b"
    "Execute transaction (decrement T_b")
    return(new T_b" for source and recipient)

The Algorithm 4 shows the pseudo-code algorithm for the transaction transfer smart contract. In the RegistryChain model, conversion to fiat is left to each organization that sell their tokens to implement the price.

4.3 Summary

In this chapter, we presented the summary of the survey findings from both the health facility mapping and the health practitioner survey. Both surveys corroborate the literature findings that Health Information Exchange (HIE) in low and middle income countries are still at the rudimentary stages.

In addition, the novel Regulated-Federated-Decentralized (RFD) framework was introduced; RFD is an enterprise blockchain-based software integration pattern. The reference implementation of RFD model, RegistryChain was modeled and discussed. RegistryChain facilitates transparent registries and repositories HIE without a central intermediary. FHIR-based ontology and data structure were also designed and presented. Finally,

the following were also mapped out and presented: the Logical and Physical architectures of RegistryChain; the algorithms for asset transfer (read and write) and transparent token economics; and tokens minting and incentives for transfers.

In the next chapter, RegistryChain, the result and evaluation metrics will be presented.

5 Evaluation and Discussion

This chapter of the Thesis is used to discuss the evaluation of the RegistryChain Ontology and Model simulation in addition to the discussion of the results with respect to the research questions. Also, the guide for using the model in practice is also discussed.

5.1 Evaluation

5.1.1 Validating Ontology

The Web Ontology Language (OWL) was chosen for modeling the RegistryChain ontology as it allows for concepts to be validated through reasoning tools. Standardized health ontology are consistent and unified framework for categorizing and describing medical terms and concepts. Established Ontologies like SNOMED-CT (Systematized Nomenclature of Medicine – Clinical Terms) and LOINC (Logical Observation Identifiers Names and Codes) provide a standardized language that ensures consistency and accuracy in the representation and exchange of some health information. While these Ontologies have existed for many decades, they are not widely used in most low income countries, largely due to skills need and also contextual complexities in adopting them. The ontology for this Thesis was designed to address this challenge using the Prodégé tool and used to adapt and extend existing ontologies. Prodégé is one of the many ontology development tools and it comes with inbuilt reasoners, with options for using plugins to add other reasoners. Evaluating the RegistryChain ontology involves structure verification, logical consistency, semantic verification, and peer review by domain experts.

The first step was using domain knowledge of the author to visually analyze the Class Hierarchy by navigating to the Classes tab, reviewing and ensuring the classes, their subclasses, and superclasses and relationships check-out as intended. For instance, the CodeSystem class has SNOMED-CT and LOINC subclass, and an instance of a Patient's record coded with these subclasses objects, with FHIR structured data properties in-

spected for health domain consistency. For logical consistency verification, the inbuilt reasoner *HermiT* was used to ensure that concepts and their relationships are validated and correct. Figure 5.1 shows a cross-section output of this validation exercise. The verification focused on sections of the ontology that was extended, with focus on the CodeSystem class.

In the standard FHIR ontology, the CodeSystem is optional and implementers are free to chose their code system of choice at implementation. This has resulted in more fragmentation and limited interoperability, with hundreds of different approach to implement these CodeSystem extensions. For this Thesis, extensions such as OutOfPocketPayment, PostDeliveryVisit, and SVD were added to the standard ontologies as extensions as seen in Figure 5.1. The HermiT reasoner was then used to reason and classify relationships of the classes and sub-classes. When the HermiT reason was run and observed for logical inconsistencies, such as unsatisfiable class, none of the classes appeared in red, showing all classes were satisfied. Similarly, a check for conflicting definitions as the reason is run returned positive as cardinality constraints and other property characteristics were properly applied and did not introduce any inconsistencies. Also, the Ontology semantics was based on the semantics of the established ontologies from the health domain used, i.e. SNOMED-CT or LOINC or ICD11. Instances of a CodeSystem class with these ontologies were observed through the individuals tab to ensure that individual instances are correctly assigned and consistent with the class definition. Finally, the ontology was peer reviewed by a public health professional, which did not involve any changes as the ontology already conformed to FHIR and other terminologies. For example, it can be seen that SNOMED-CT is both a poly-hierarchical CodeSystem and a ClinicalTerminiologyRegistry, while ICD11 and LOINC are both uni-heirarchical CodeSystem and ClinicalTerminologyRegistry. This makes it easy for users to quickly understand which implementation is most optimal. The implication of validating this ontology is that extensions are free of relationship and concept errors. Our custom ontology was found to be both consistent extension of current FHIR and SNOMED-CT ontology and also error free.

5.1.2 Simulation

Zhuang et al. already conducted an Ethereum simulation by passing Patient data between nodes [50]. They simulated 1.5 million data request transactions for months continuously, where they received 100% approval. They noted that physicians received on-chain data within 20.398 seconds. Similarly, Pratap et al. simulated and recorded 1000 transactions for different data stored on peers in docker container [51]. They recorded the average

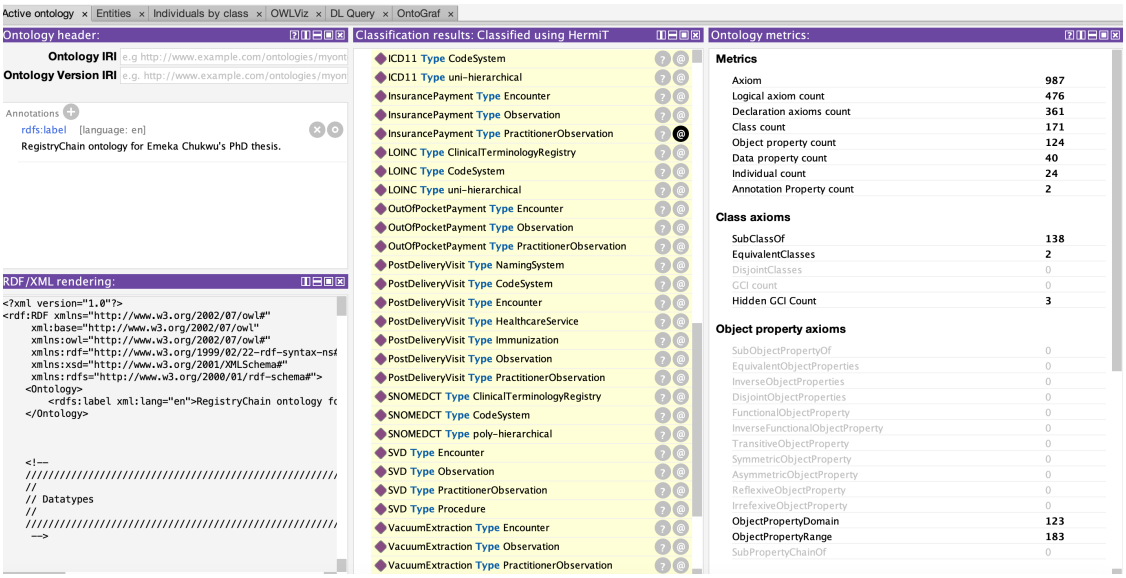


Figure 5.1: Cross-section of output of validated Ontology (Prodégé generated)

latency of between 26.22 seconds and 36.67 seconds for opening a transaction and between 0.12 and 7.62 seconds for querying a transaction.

The BPMN in Figure 4.8 details the process for a Patient, Health facility, and Regulator in a complete flow of the transaction. For performance testing, please note that the Patients are registered as part of the Hyperledger Fabric user enrollment operation, which is also the step where an X.509 certificate with the PKI is generated for the user. Organizations can join the network or call special contracts, which generate a needed token. The evaluation of RegistryChain focuses on RAM, CPU, network, and disk-storage utilization.

The tests for the model were setup on a host-based setup using Ubuntu version 20.0, on HP Elitebook 2GHz, and 8GB memory. Each component was setup as a docker container in the host. Based on the structure of data passed in the health sector, two data types were identified as candidates for simulation and performance evaluation. The first is discrete data only type (summary or aggregate data, formatted as integer type). The second is the individual patient-level data type consisting of other data types, including strings. For each data type, three functions were created - one to create a record, a second to query one record, and the last to query all records. These simulate the main operations that a typical blockchain system will be executing. In all, there are six functions deployed for testing and evaluating this RegistryChain.

5.1.2.1 Discrete data

The discrete-aggregate data like the number of practitioners, their roles, and services for a given health facility was randomly generated in FHIR format. Figure 5.2 illustrates aggregate information representing the most used type of data for reporting is shared for personnel information for a particular health facility. Each record represents an update of the health facility record. Assuming the health facility personnel information, certification, role, and other characteristics change often. The Random Access Memory utilization was also measured for each of the smart contract operations. Each of the *createonerecord*, *queryonerecord*, and *queryallrecords* use average of 65MB except for the certificate authorities and the orderer that uses 11MB and 13MB for these operations, respectively. See as illustrated in Figure 5.3.

The chart in Figure 5.4 captures the caliper simulation output documenting network bandwidth utilization of the docker containers and the amount of data stored or read for each container for each aggregate data smart contract operation.

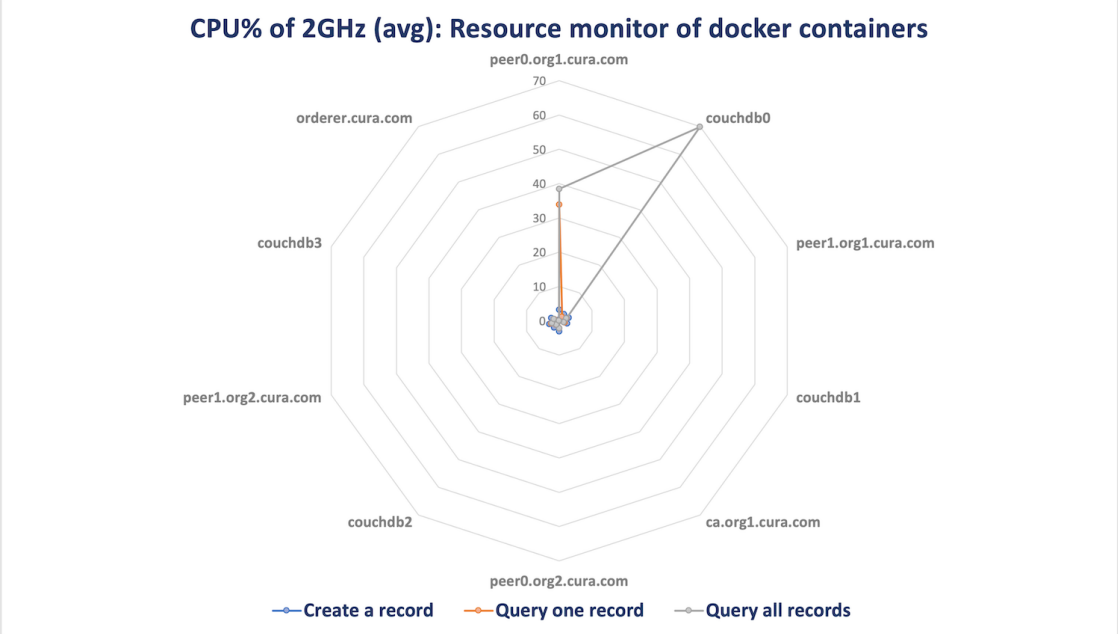


Figure 5.2: Aggregate data processing CPU utilization

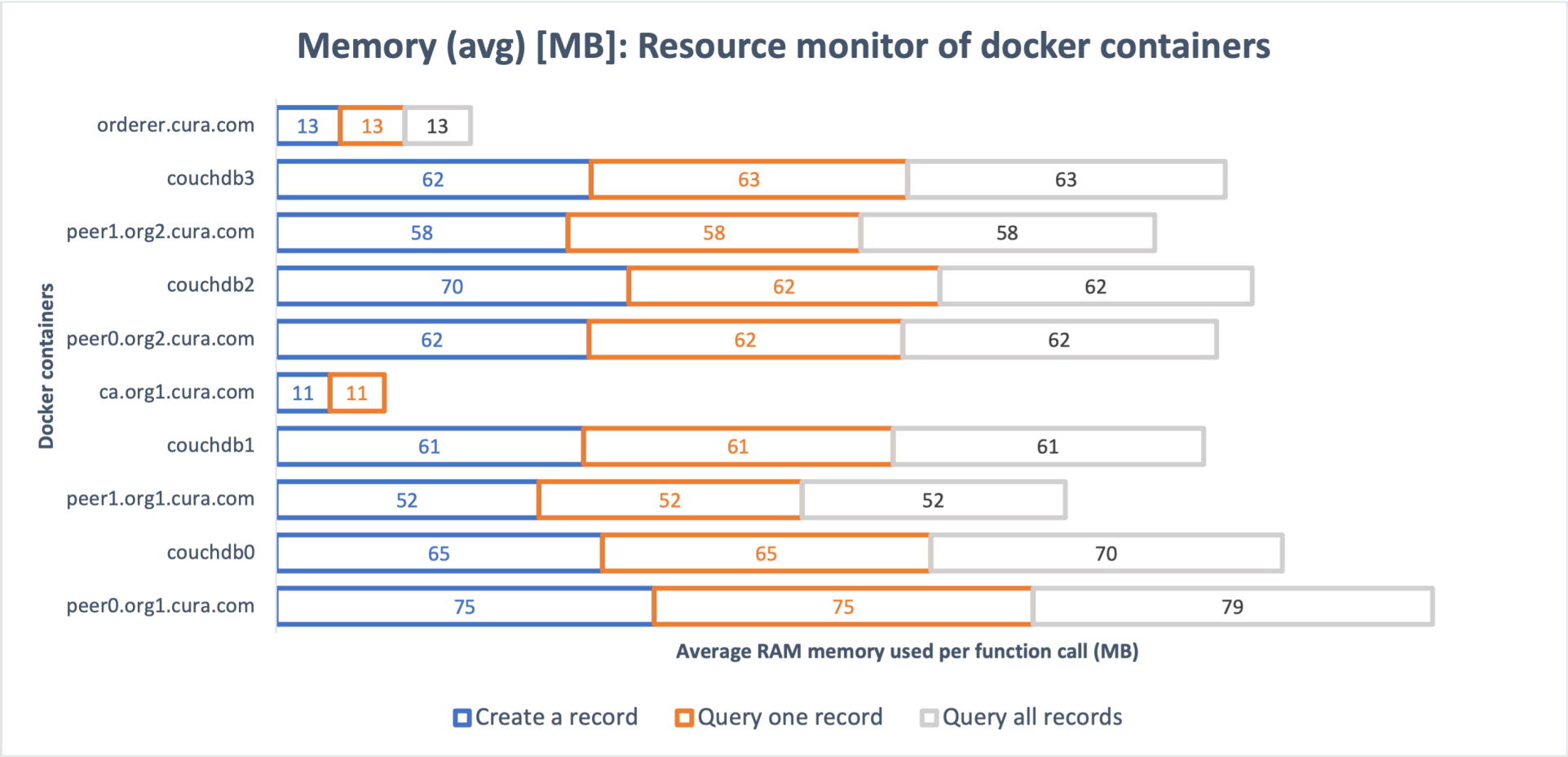


Figure 5.3: Aggregate data RAM utilization

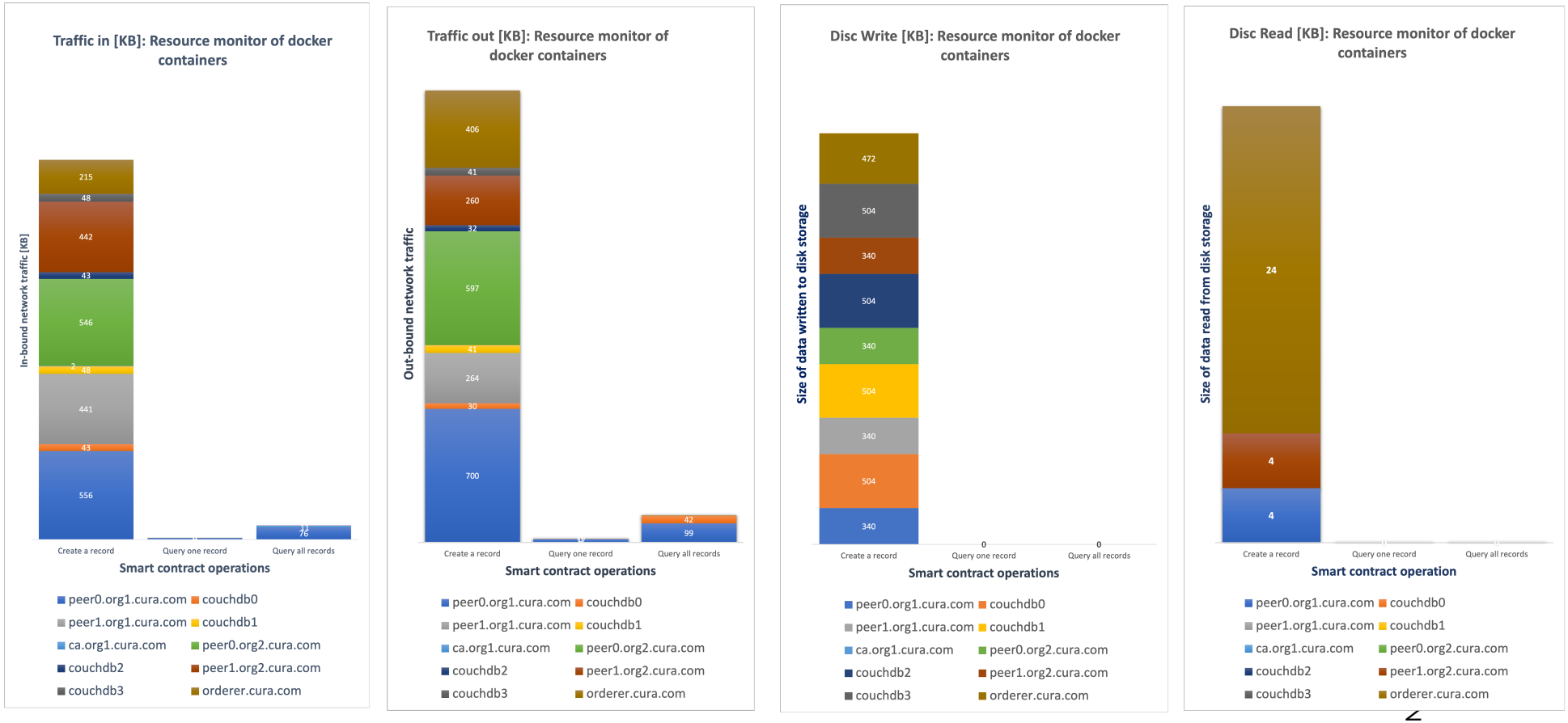


Figure 5.4: The aggregate data network and disk storage utilization

5.1.2.2 Patient-level data

For the patient level information, data from actual US cancer patients were obtained from US National Cancer Institute $Surveillance, Epidemiology, and End Results (SEER)$ project [275]. The data of 1.048 million patients (both dead and alive were used), and to get a manageable number, the living patients were filtered for inclusion. The data element $Vitalstatusrecode(studycutoffused)$ was used to filter in the 282,252 alive patients and exclude the rest. The data was cleaned to reduce the indicators from 23 to 10 indicators to eliminate duplicate data record indicators. As a script for calling, one created record from real patient data was called and saved to the blockchain RegistryChain network. The CPU utilization, inbound and outbound network utilization, memory use, and disk read and write for each docker container were also measured. See Figure 5.5 for CPU utilization.

Similarly, the memory utilized for each container is shown in Figure 5.6. Also, the memory usage for each container was fairly constant at 60MB. Also, the memory usage for the certificate authority and the orderer were also negligible (at 10MB or less) except for $readall$ operation on the orderer when memory jumped to 40MB. See this in Figure 5.6. Similarly, Figure 5.7 illustrates the network utilization when patient-level data is created and queried for each docker container representing blockchain nodes.

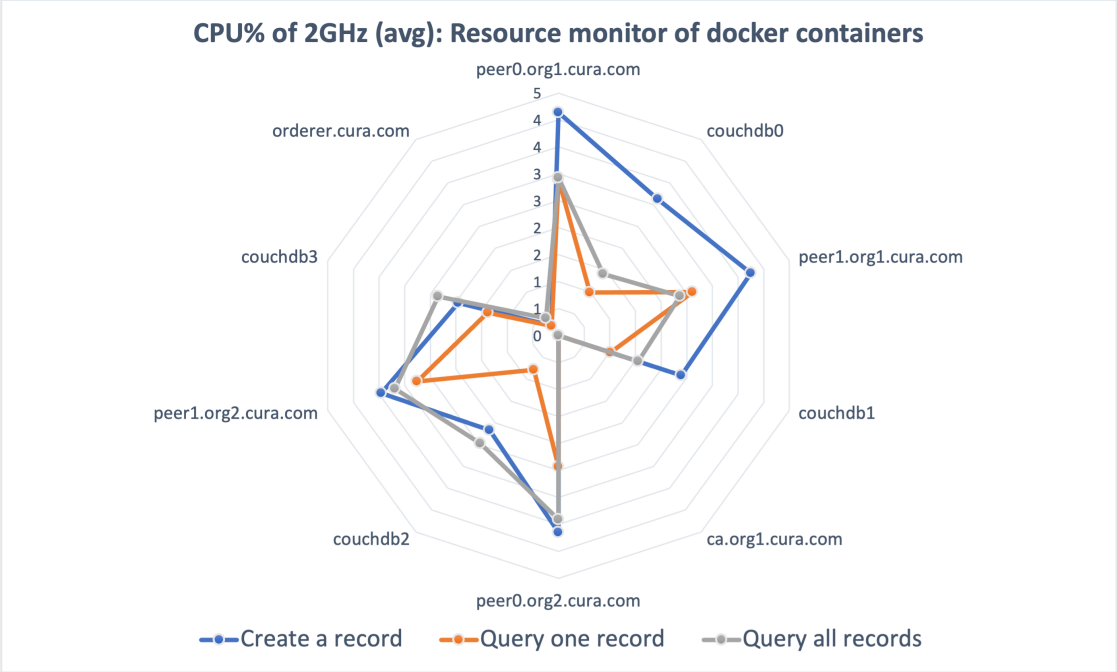


Figure 5.5: Patient-level data processing CPU utilization

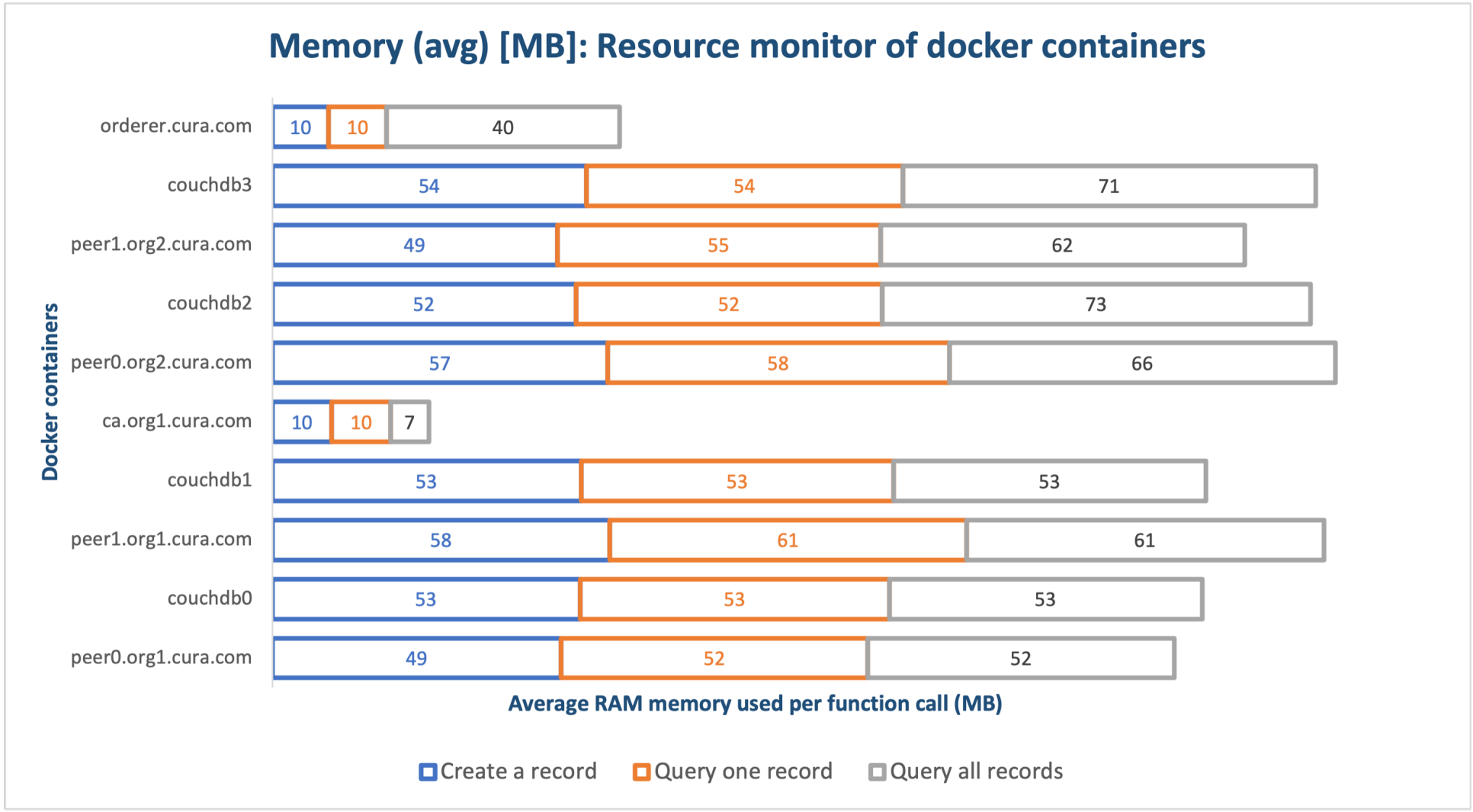


Figure 5.6: Patient-level data RAM utilization

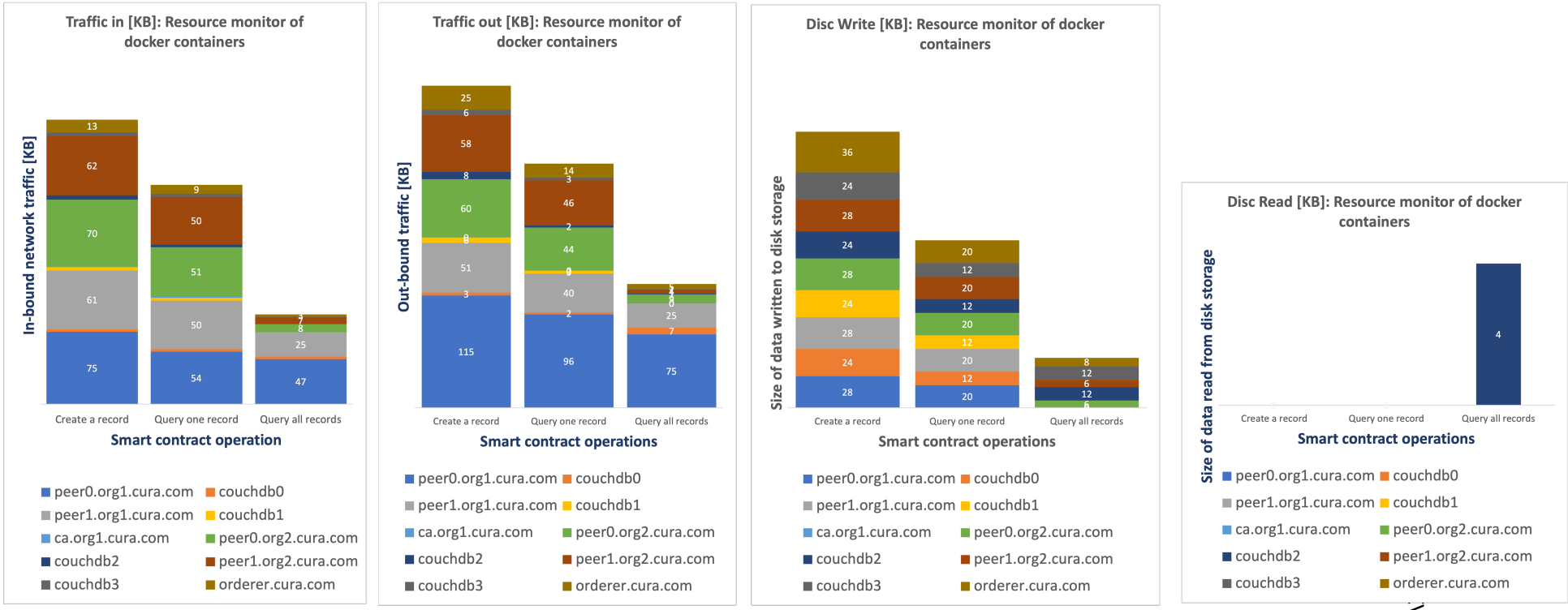


Figure 5.7: The Patient-level data network and disk storage utilization

5.1.3 FHIR structured Terminology data

The CodeSystem FHIR resource structures terminologies in the healthcare data processing. As discussed earlier, one example of a codeSystem is the SNOMED-CT terminology set. The CodeSystem section of the RegistryChain proposed ontology groups CodeSystems into two broad categories - established clinical terminologies and NamingSystem (also an FHIR resource), see Figure 5.8. The established ontologies are the tested and widely used schemes like SNOMED-CT, ICD11, LOINC, and the like. In addition to this established terminology semantic standards, there are custom and less globally established standards used from one locality to another. Things like drug lists are issued from time to time by government health insurance regulators within each jurisdiction. Others are list of insurance services or a list of health facilities and their coding, to name a few. This thesis recommends structuring these local standards as a Naming System resource.

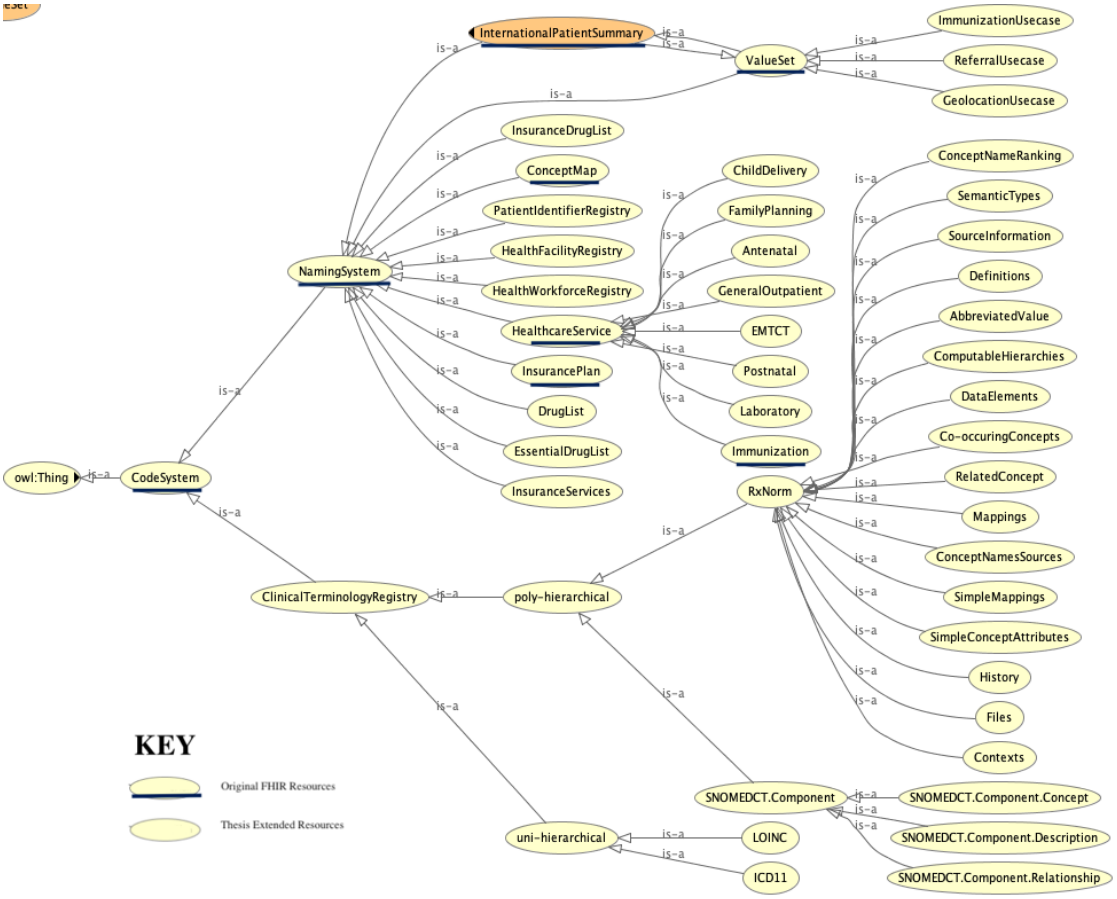


Figure 5.8: Proposed CodeSystem Ontology

A simple example of a terminology resource is shown in Figure 5.8

CodeSystem.json > ...

```json
{
  "resourceType": "CodeSystem",
  "id": "summary",
  "text": {
    "status": "generated",
    "div": "<div xmlns=\"http://www.w3.org/1999/xhtml\">\n  user define text here </div>"
  },
  "url": "http://hl7.org/fhir/CodeSystem/summary",
  "version": "4.0.1",
  "name": "Code system summary example for body sites",
  "status": "draft",
  "experimental": true,
  "publisher": "HL7 International",
  "contact": [
    {
      "name": "FHIR project team",
      "telecom": [
        {
          "system": "url",
          "value": "http://hl7.org/fhir"
        }
      ]
    }
  ],
  "description": "This is an example code system summary for codes for body site.",
  "useContext": [
    {
      "code": {
        "system": "http://example.org/CodeSystem/contexttype",
        "code": "species"
      },
      "valueCodeableConcept": {
        "coding": [
          {
            "system": "http://snomed.info/sct",
            "code": "337915000",
            "display": "Homo sapiens (organism)"
          }
        ]
      }
    }
  ],
  "caseSensitive": true,
  "content": "not-present",
  "count": 92
}
```

Figure 5.9: Example FHIR CodeSystem resource adapted from FHIR Rev 4.3

The FHIR-structured terminology data in JavaScript Object Notation (JSON) is a standard format for the update to each practitioner's information, their roles, or licensing information. Figure 5.9 illustrates a simple example FHIR resource for CodeSystem but shown in Figure 5.10 is the complete CodeSystem resource data structure. The data structure illustrates the extent of complexity possible from this data structure. Many optional data fields in the CodeSystem resource may be mandatory in certain established terminology dataset systems.

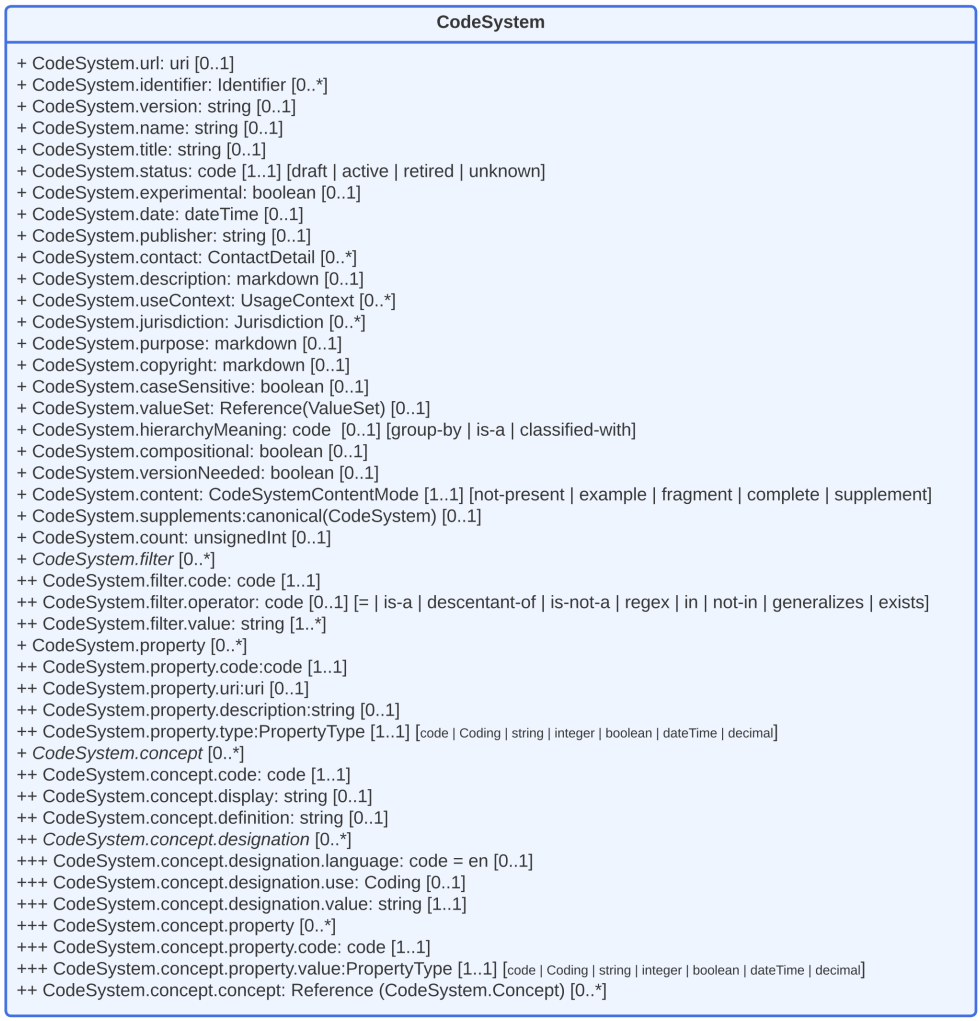


Figure 5.10: The CodeSystem FHIR Resource data structure

A sample FHIR-structured terminology information was prepared and stored on the blockchain. Each record represents an update of a terminology record. Given that this information contained larger-sized data, 10,000 records were created, and records were searched up

by ID. First, querying for a record within the first ten records (representing records in front) and querying for a record between 5000 and 5050 in the database (representing records in the middle). Finally, one record at the end between 9050 and 10,000 was queried. For all queries, the processor speed utilization, the Random Access Memory utilization, and Network utilization was recorded for all peers, orderers, and certificate authorities.

The chart in Figure 5.11 shows the average percentage of the 2.0GHz processor used when conducting the search for each query. As can be seen, while there is higher processor speed utilization by each peer for the middle operation, the difference cannot be said to be significant. One thing that standards out are that the orderers and certificate authorities of the three organizations consume negligible processing speed (between 2.7% and 3.5% of CPU speed).

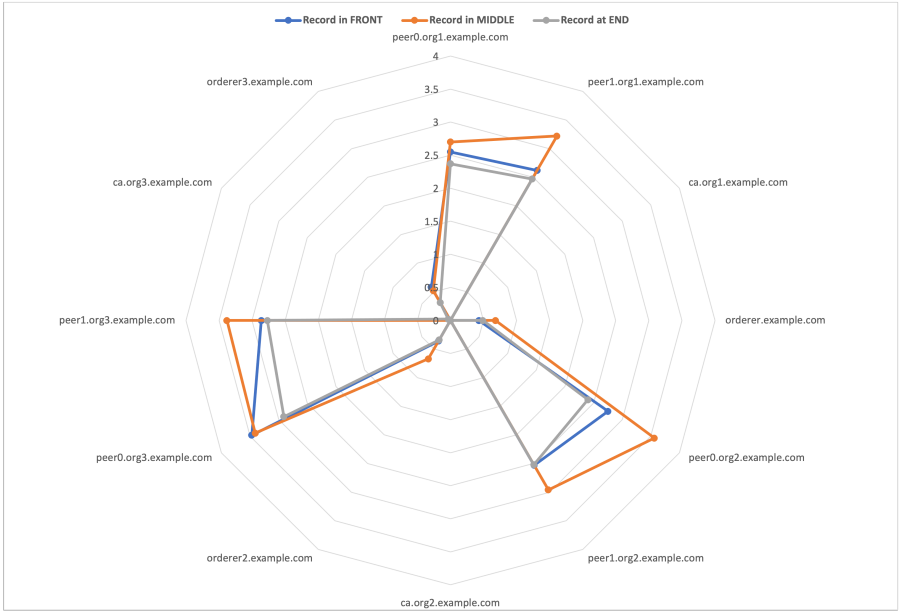


Figure 5.11: Blockchain simulation (Processing speed)

For each of the *query*, each peer use average of 65MB except for the certificate authorities and orderers that use significantly lower memory than the peers. However, the peer that initiated the query, in our case the peer0.org1.example.com consumes significantly more memory as shown in Figure 5.12. And this is the case for each record query. Similarly, the certificate authority and the orderer nodes consumed an insignificant amount of memory.

The chart in Figures 5.13 and 5.14 captures the Hyperledger caliper simulation output

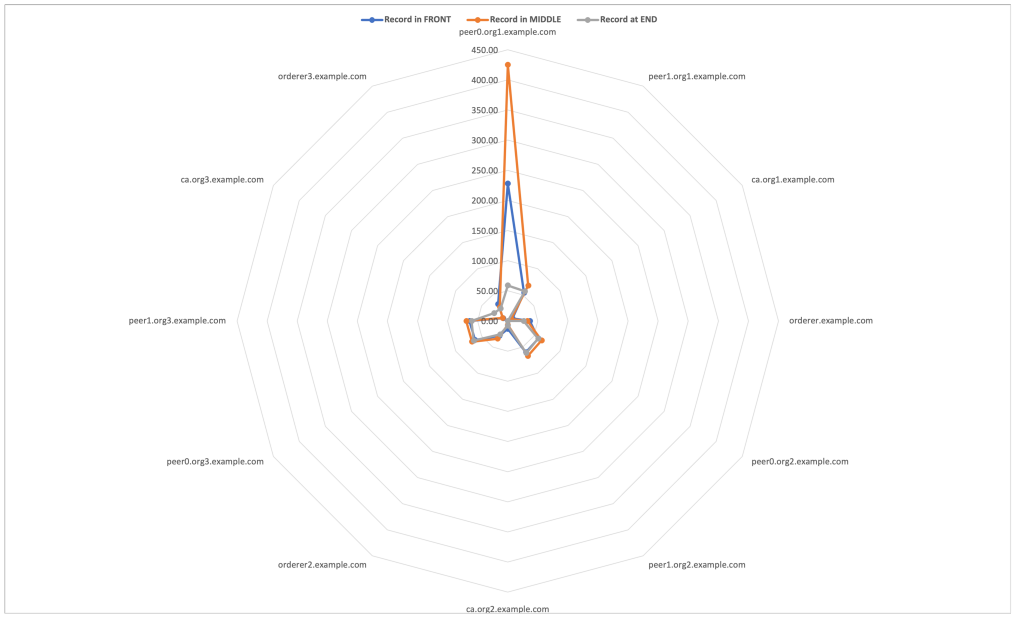


Figure 5.12: Blockchain simulation (Memory used)

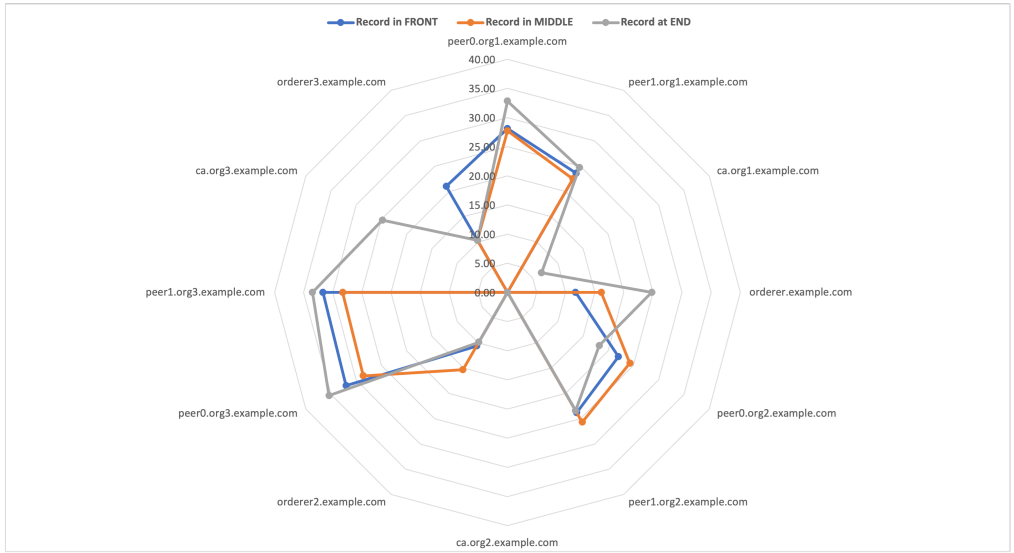


Figure 5.13: Blockchain simulation (Network in)

documenting network bandwidth utilization of the docker containers for in-bound and out-bound traffic each node by the three organizations. The network utilization for each organization's nodes ranges between 0KB and a maximum of 35KB. This means the network bandwidth requirement to run a node with this model is within what most low-resource environments can handle.

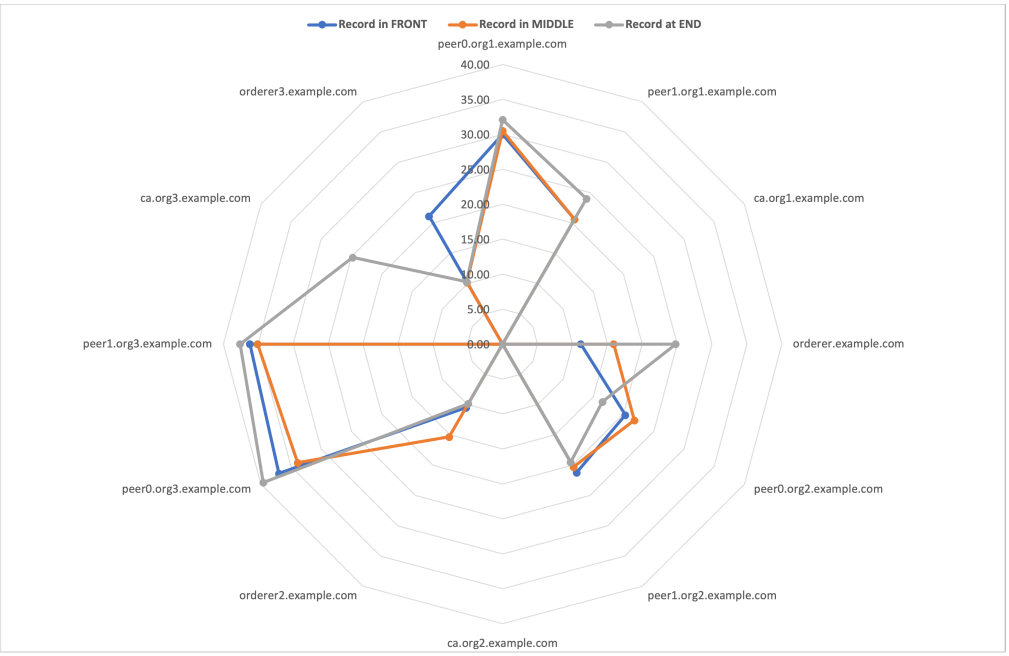


Figure 5.14: Blockchain simulation (Network out KB)

5.2 Discussion

This Thesis has designed and presented a novel decentralized framework that can be used for shared data and token custodianship. It also presented the formal model and data structures to make this possible. Both hypothesis for this Thesis were supported.

5.2.1 A decentralized and oracle-friendly software integration pattern will facilitate stakeholder shared data custodianship and economic value

This hypothesis is supported by addressing two research questions posed at the beginning of this Thesis. First, that only a decentralized framework like blockchain can support information sharing without a central authority. Then that it is possible to leverage the token feature of blockchain to design a model token economy for healthcare data sharing. Each of these research questions and their contribution are discussed in further details next.

5.2.1.1 Decentralized Information Sharing without central authority

The research question of which Health Information Exchange (HIE) framework will facilitate will facilitate HIE with shared data custodianship is address through the literature. The literature show that blockchain-facilitated information sharing is the only integration pattern without a central authority. As part of this Thesis, the systematic literature review and analysis of the application of blockchain in healthcare show that many experimentation and prototypes, few implementation across multiple health use case of real world applications of blockchain in healthcare exist [1]. Implementing blockchain frameworks comes with significant cost implications as our earlier research show that some blockchain networks can be expensive to store basic data, especially the proof of work based blockchain models. As a result, enterprise modeled blockchain model like RFD is most appropriate for adequate cost. Also, a blockchain based information storage allow for shared storage and ownership of data stored on the ledger. This therefore supports the hypothesis that a decentralized integration pattern will support share data custodianship.

5.2.1.2 Economic value sharing without a central authority

The research question on the optimal model that will facilitate transparent and shared data economic value did not find existing shared economic-value frameworks in literature. It is safe to assume that traditional central authorities responsible for managing information sharing workflow in a health (or other domain) enterprise is also responsible for economic value sharing. This is particularly so as the few health information exchange use cases are centrally managed either as one organization owned API-based system, like in Denmark [70] or a multi-organization (consortium) owned central API-based system like HealthEConnections in the US [83]. Also, very little has been published on this concept of shared economic value in healthcare. Yet lessons from other sectors show that tokens are crucial for value extraction from an enterprise including in healthcare. In healthcare, service value remain important, as for instance, US pay service providers to provide digital services [68]. In Low and Middle Income Countries (LMIC), non-blockchain tokens have been implemented for health services in both developing and developed countries [67], but largely by a central authority, limiting its utility, scalability, and ability of other competitors to participate in shared value. Shared token is also a concept already popular with cryptocurrencies, but they are not enterprise mature and to our, many workflows for healthcare has not been defined or published. Traditionally, the central organization is often the one responsible for token generation and management, with little or no transparency for stakeholders not involved on the platform ownership. A shared token will ad-

vance transparency and sense of shared ownership, which in turn can drive up adoption, and thus supporting the hypothesis. In LMIC countries, breaking down these challenges mean reducing barrier to *entry*. Shared-tokens as used in blockchain based systems if implemented in the health sector, means that transactions will not have to wait for audits or scheduled reconciliation rounds, but can be seen and possibly withdrawn in realtime. this can greatly increase trust in the equity of economic value and facilitate greater collaboration.

This Thesis defined a novel workflow and token management workflow, metric, and algorithms. These are developments that are no where featured in any published literature. The ERC20 addressed a workflow and corresponding token for *organization-join token minting*, *integrity token award*, *asset commit/read token minting*, and *Token asset transfer*. The algorithms and the sequence diagram were clearly defined and presented. The Business Process Modeling Notation (BPMN) for the workflow was also developed in context of typical LMIC to reduce the barrier to *entry* for both health-business or IT-teams designing and implementing interoperability solutions.

5.2.1.3 Stakeholder shared data custodianship and Oracles

In the health sector, there are many Oracles, some public sector Oracles, and others are private sector Oracles; though not specifically called Oracles in the health sector, they represent single source of truth for certain healthcare registry data points [35]. A typical example is a Terminology service custodian and provider may be mandated by law to manage and make available terminology service (example say a custodian of SNOMED CT) in a jurisdiction in LMIC [12]. Such a custodian is an Oracle for the terminology database, they make updates and want to transparently manage requests for changes and make changes happen and push changes back to stakeholders. Without a blockchain-based model as we have proposed, such changes will be opaque to most stakeholders and only available to the central authority, and can result in further resistance to adoption. My conclusion is that this transparency in transaction can increase uptake especially as perception grows. And to my knowledge and based on my research, this is the first proposition of such shared service management. Also, a patient identity management databased for a Master Patient Index (MPI) which could be a multi-sector national ID or a health sector ID management system will generally be an Oracle, so also is a drug formula database often called *"registries"* or *"shared records"* [61]. These are Oracles whose source of truths come from the respective organizations, but made available to the different implementers. This will mean that the different stakeholders will manage their respective EMR data while

accessing and viewing how Oracle change decisions are made, or receiving Oracle data update and ingesting them into their system without complex data sharing agreements. A channel can be created for certain Oracles, so they can continue to perform their Oracle functions, while participating in the blockchain network. This is a key contribution of this Thesis and also help support the hypothesis that a decentralized, oracle-friendly software integration pattern like I have proposed will facilitate shared data custodianship and shared economic value from the data.

5.2.2 Standardized healthcare registries and repositories on permissioned shared-ledger will facilitate Health Information Exchange (HIE) without an intermediation trusted party.

This Thesis also supported the hypothesis that datasets for shared registries or repositories need to be standardized starting from established health domain ontologies. It is already established that a decentralized health information sharing framework without a central authority was feasible and is the integration pattern that allow share data custodianship. Data sharing in healthcare require standardization of datasets and terminologies, and the RegistryChain ontology support this hypothesis.

5.2.2.1 Global State of Health Information Exchange (HIE)

Several enterprise software integration (and information sharing) patterns exist, from information portals, data replication , shared business solutions, distributed business process, Service Oriented Architecture, and Business-to-Business. Yet, information sharing, particularly health information sharing remain an uphill task. In healthcare, concerns for how to integrate information from multiple systems is particularly important as a patient will often encounter (leaving information crumbs) many physician and healthcare practitioners in the course of their care continuum. As these problem manifest in form of data sharing or process sharing by multiple stakeholders - Regulators, Patients, Vendors, Private Sector, and Health System [6]. From the literature review, existing health information exchange models has been largely Centralized with a few cases of Federated and Patient centered. Decentralized models based on blockchain are emerging and most publications have not provided details. Even in economies with mature HIE, Bernstam et. al. found that intra-EHR information was possible only 68% of the time, and inter-vendor information sharing was as low as 22% of the time [276].

The literature as in Chapter two show many deployments of HIE in the US and EU, and few experimental deployments in Africa and other developing world. The HIE components such as mechanisms for uniquely identifying patients, providers, institutions, devices, and software. Interventions for health information sharing in the UK and EU has been setup since 1994. In the US, it started with Health Information Technology for Economic and Clinical Health (HITECH) with several Regional Health Information Organization (RHIO)s augmented by State Health Information Exchange (HIE)s. On the African Continent, there are many experimentation and HIE simulations. Asia region has the same story as the Africa region, especially from published literature. Barriers identified to information sharing in health sector include Technical, Motivational, Economic, Political, Legal, and Ethical Barriers.

5.2.2.2 Implication for Low and Middle Income Countries (LMIC)

The findings from health facility mapping show that none of the health facilities surveyed in Sierra Leone electronically shared clinical information, we have already published this work [2]. At the same time, only 23% of hospitals share aggregate (summary formatted information). This means that no standard or terminology is defined or used as no clinical information is shared. Similarly, a small sample of health workers in Ebonyi state Nigeria were surveyed at how referral is managed at their health institutions. In addition, the experiential knowledge of the researcher show that there are limited EMR vendors using standards or sharing clinical information in Nigeria. The survey of health workers on the other hand as already published here is important [3], and the author acknowledge the potential of selection bias. Notably, the providers were mostly from health facilities in the state capital, Abakaliki, which has better computing infrastructure. This bias is considered insignificant as the measured variation in referral datasets amongst providers's responses was significantly different. If the variation was uniform, then the bias would have mattered, but the variation was significant and different, thus reinforcing our earlier hypothesis that information sharing happen at a limited scale in LMICs. Also, given that each provider was from different health facilities, this can be considered representative. This represents a confidence interval of 20 at a confidence level of 95%. In addition, the three different referral forms used in the state for pregnant women referral tracks is also an indication of the variance. This show that information sharing does not happen, and that data used for sharing health information are also not standardized.

Health Information Exchange (HIE) can help solve many of the health systems challenges facing Low and Middle Income Countries (LMIC). Notably, care coordination amongst

different healthcare providers will be easier, resulting in better quality healthcare for the Patient. This improved coordination as a result of optimized HIE will mean better accountability and better resource utilization, reducing waste and errors overall. In LMICs, resources (notably human and material) are extremely limited and any improvement in efficiency can result in significant savings [277]. While best practice is to base public health decision on trusted healthcare data, for most LMICs, public health data have been of poor quality resulting in reliance on annual (and sometimes every 5 years) surveys [278].

5.2.2.3 Low computing resource capabilities

The simulation result has shown that even for FHIR-structured terminology data, it can be saved and retrieved on the permissioned-blockchain with little processing speed, memory utilization, and network overheads. The feasibility shows that this is possible in low- and middle-income countries with limited resources. The simulation result shows that in a low-resource environment where electricity, network, and computing infrastructure are limited in RAM and CPU processing speed, the model and data for both discrete and patient-level is feasible. For instance, the resource needs is so low that it can run-off a raspberry pi style computing hardware. The enterprise style blockchain just needs signatures and not proof-of-work algorithms, something used in other domains like logistics and finance sector, but still being experimented in health sector. Due to the high energy needs of proof-of-work consensus algorithms, most low income countries cannot participate in such sharing economies. A signature based model is light weight in terms of processing, storage, and network needs, thus easy to find applicability.

These LMIC countries are notorious for poor computing infrastructures such as low network bandwidth, limited power supply and low grade computing hardware (processing power). Something that was established in our earlier published paper [279]. The simulation results of RegistryChain show that utilization of these resources for bandwidth, processing speed, and storage space for the model proposed in this Thesis is adequate for the limited resource environments like typical Low and Middle Income Countries (LMIC)s. This means that leveraging any of the established ontologies or the RegistryChain model ontology can speed up health information sharing in a typical LMIC, thus supporting the hypotheses. The implication is that low resource environments can use low processing speed hardware for a lower cost.

5.2.2.4 Healthcare standards used for Semantics and Syntactic HIE

The standards used in healthcare are broadly grouped into those for the syntax of the language formatted in JSON and XML formats and those for the semantics of medical terms [11]. Both have been presented and discussed as part of the proposed model. Based on my systematic literature search and analysis, HL7 FHIR was found to be the most widely used healthcare standard for syntax and structuring of healthcare data [1]. The standard was subsequently used for the RegistryChain Ontology. Similarly, there was no clear semantic standard mostly used in health sector. Different semantic (or terminology) standards are strong for different health areas. Though the most comprehensive of all healthcare terminology standards is the SNOMED-CT, and it is not widely used for technical and cost reasons. Using FHIR for syntax and other terminology (semantic) standards like SNOMED-CT will greatly improve standardization in low income environments as FHIR is already based on widely used RESTful API framework.

5.2.2.5 Optimal Architecture for non-centralized data storage and data-processing

It was earlier established in section 5.2.1.1 that the integration architecture that permit information custodianship that transparently allow regulatory (oracle) participation will have be decentralized architecture, and it lends itself well to transparent audit. Similarly, data custodianship, and its implication for healthcare Oracles has been discussed in section 5.2.1.3. A blockchain based system like RegistryChain allows for information sharing on multiple nodes based on any defined architecture. Similarly, the information can be served to requesting clients by designated nodes within an organization, contributing RAM, network, and processing resources as needed. A sharing economy based on RegistryChain model can then be used to implement incentives for the model of shared data custodianship and shared data-processing.

5.3 Guide to utilizing RegistryChain in a health system

A typical health system will have many user concerns. Stakeholders will understand and prioritize appropriate healthcare use cases to demonstrate (or pilot) transparent data and economic value exchange and collaboration. The steps will involve digital readiness assessment, Stakeholder identification, and determination of Network models.

5.3.1 The business requirements catalog

Any health system that wants to integrate digital operations will need to first assess the state of digital intervention in the ecosystem. This assessment will detail current (AS-IS) architecture, security, critical resources, the risks, amongst others. The assessment will detail available network, computing, and electricity infrastructure. The assessment should also catalog disaster AS-IS and TO-BE disaster recovery processes. The output of this stage will be architectural artifacts, mostly in UML formats. In addition, a risk register for implementing a decentralized and transparent system such as RFD should be developed. The business catalog of a use case will determine the data needs, stakeholders, and their concerns.

5.3.2 Stake, policy, standards, and plan

The AS-IS and TO-BE output of the current business process for identified stakeholders and select use cases will feed into the network stake and plan. Consortium members will have to agree on stake sharing, endorsement policy, and data commit structure. The default stake for RegistryChain is an equal stake for all stakeholders. Similarly, the default RegistryChain endorsement policy is 'MAJORITY' endorsement for chain code and transaction commits. Most use cases can use the proposed default configuration. Similarly, the proposed model used FHIR for use case data structuring. The RegistryChain model also used SNOMED CT poly-hierarchical model. The data structure for the health system selected use case can be determined from stakeholder data needs and synced to appropriate FHIR resources. All of these are implemented as part of the smart contract implemented for their specific channel(s). Though the thesis illustrated the smart contract implementation using Javascript based on node.js, it can be implemented using any of Java or Go.

5.3.3 Network setup, pilot and scale

The network can be setup as described above. The network structure above proposes the peer, orderer, and CA nodes to be on separate containers on one machine. A minimum of six months is recommended for an initial pilot in a health setting before scaling up.

5.4 Comparison to similar frameworks

Many different aspects of Health Information Exchange (HIE) have been discussed like the health system problem solved, the technical standards used, and how individuals are

identified. Others are data dictionaries adopted, the governance architecture, the Certificate Authority (CA) model, type of information exchanged, security & privacy management strategies, performance in Transactions Per Second (TPS), and approach to token and incentive management. The main distinguishing factor between the proposed framework RegistryChain and existing work is the use of the shared token for healthcare data management while managing identities using multiple certificate authorities. RegistryChain also integrates standards both for Semantics (Terminology) management, example SNOMED CT, and Syntax, example FHIR. The framework TPS is the highest for all framework that supports Authoritative Registries with trusted shared contents (Oracle), and decentralization.

The Table 5.1 illustrates some of the differences between RegistryChain and existing frameworks.

Table 5.1: Comparing Health Information Exchange (HIE) Frameworks

Paper	**Governance**	**Oracle**	**Syntax**	**Semantics**	**CA identity**	**TPS**	**Token**
[50]	Decentralized	Yes	Yes	No	PoW	4	No
[106]	Centralized	Yes	Yes	Yes	No	nil	No
[280]	Centralized	No	No	No	One-CA	nil	No
[109]	Federated	Yes	No	No	One-CA	1782	No
[107]	Centralized	Yes	Yes	No	One-CA	10	No
[108]	Centralized	Yes	Yes	No	One-CA	120	No
[112]	Federated	Yes	Yes	Yes	One-CA	nil	No
[120]	Decentralized	No	No	No	PoW	nil	No
[122]	Decentralized	Yes	No	No	PoW	nil	No
[128]	Decentralized	No	No	No	PoW	46	No
[129]	Decentralized	No	No	Yes	PoW	170	No
[130]	Decentralized	No	No	No	One-CA	nil	No
RegistryChain	Decentralized	Yes	Yes	Yes	Multi-CA	142	Yes

The Framework SeSPHR, proposes a methodology for secure sharing of PHR in the cloud describing how data is stored in an encrypted format and accessible from a trusted server [280]. It did not discuss the type of data, the standards used or how multi-parties can manage access. The framework assumes a centralized architecture model just like [106] and [107]. Osebe et al. proposed a Digital Health Wallet for the management of health data, though the token algorithmic detail was not discussed [281]. Also, the EPMS was proposed for large-scale patient identity matching that leverages fuzzy logic but was designed for a central governance environment [282]. Cross-organization access control

for EHR was proposed and simulated by Ma et al., and RegistryChain leveraged the idea proposed there for the Multi-CA model [283].

5.5 Summary

In this chapter, Key contributions of this research has been outlined according to the research questions and hypothesis. The implication of the global state of art of Health Information Exchange (HIE) findings from literature and health facility and health practitioner surveys were discussed. The RegistryChain ontology based on FHIR attribute was validation and discussed. The simulation of smart contract operations followed this. The chapter concluded with a guide to implementing RegistryChain in a health system. The current global state of the art of HIE was first discussed, with focus on Low and Middle Income Countries (LMIC). Then the implication of the surveys of health facilities and health workers in Sierra Leone and Nigeria were discussed and the implication of the survey findings and how it complements the existing literature on health information standards and sharing. The implication of the results with respect to the HIE and integration patterns was discussed. Also, the model of how multiple parties can transparently share economic value arising from sharing healthcare data was highlighted. The proposed framework is compared to the existing frameworks to support many real-life HIE problems. Finally, how this impacts how Oracles and regulators operate in a decentralized environment is also discussed.

The next chapter will summarize this work's main contributions, industrial significance, the implication for the health system, and its limitations. Future work was also discussed.

6 Conclusions

The Thesis aimed to design an optimal framework for multi-stakeholder Health Information Exchange (HIE) in resource-limited environments. From the systematic literature search, the global state of the art in HIE was itemized and detailed, and it showed that health information sharing still face significant barriers especially in low computing resource environments and Low and Middle Income Countries (LMIC)s. Some of the barriers to healthcare data sharing, use, and evidence of the effectiveness of HIE, and data standards were highlighted. The different software integration patterns and HIE architecture models by governance model were aggregated. The different Patient identification mechanisms, the different Certificate Authority (CA) options for healthcare organization identity management, communication standards, and vocabularies were presented. A national health facility survey was conducted in 72 health facilities in all 13 districts to gather information-sharing requirements for our model show information is still not shared. The survey found that individualized health information is not being shared, though the aggregated summary is shared. A health practitioner survey was also conducted for the same purpose and found that referral forms vary widely.

The Regulated-Federated-Decentralized (RFD) framework, a novel blockchain-based software integration architecture, was designed and presented. The Regulated-Federated-Decentralized (RFD) framework combines the features of Federated and Decentralized HIE architectures, keeping their best feature which hitherto was mutually exclusive - i.e., regulation and decentralized. The RFD's reference implementation, RegistryChain, demonstrates how multiple blockchain nodes representing healthcare stakeholder organizations can keep their generated data or their data-processing custody while participating in HIE. Also presented is a novel token economy framework and algorithm that facilitates network-stake ownership, node-integrity facilitation, data-size influenced transaction-fee management, and function call fee.

The hypothesis that a decentralized and oracle-friendly software integration pattern will

facilitate share data-custodianship and shared-economic value was supported through different research questions. The Thesis also show that standards for health datasets for repositories and registries (managed by oracles) are essential for information integration and interoperability. Therefore, the main contributions of the Thesis include:

- The current global state of the art of Health Information Exchange (HIE);
- RegistryChain HIE framework that:
  - make available an LMIC ontology that includes both uni-hierarchical and poly-hierarchical terminology sets;
  - helps facilitate multi-stakeholder registry information change management with no intermediary;
  - is privacy-preserving and transparently shared health data and shared economic value like token amongst stakeholders;
  - supports communication where communicating parties do not have to all be online at runtime; and
  - facilitates data and service integrity, quality, and ease of audit.

The literature search identifies FHIR as the widely used healthcare syntactic standard. Similarly, the model identified SNOMED CT as the healthcare terminology (semantics) standard with the largest number and relationships of clinical concepts. The RegistryChain ontology, whose shared data structure is based on four FHIR bundles, was then modeled. The use case of FHIR structured IPS was set up using the Hyperledger Fabric blockchain network with multi-CA architecture. The smart contracts were implemented in Hyperledger fabric using the Node.js environment. RegistryChain was simulated to demonstrate using aggregate data used in a typical public health data interchange and real patient-level data from the US SEER database and was found to perform well for the three reference smart contract operations. This work represents a turning point in transparent multi-stakeholder, multi-vendor data, and economic value sharing without the need for a central authority while meeting privacy, security, and regulatory compliance. RegistryChain implementation of RFD will go a long way in facilitating shared custodianship, shared trust, and shared economic value while participating in HIE. Also, From the literature available, this is the first system that represents how tokens can be used in a Health Information Exchange (HIE) network.

6.1 Industrial significance

The survey shows limited evidence of HIE, particularly in LMIC. RegistryChain simulation shows low bandwidth, low memory use, and low disk-space utilization, meaning disconnected health facilities can use low-resource hardware to run docker-deployed containers. This will help facilitate a transparent, shared economic model, fast-track information sharing, and help drive seamless integration in health systems globally, especially in low-income countries. One area needing such integration is national vaccine roll-out information sharing. This model will ensure that software vendors (and database custodians) can keep their data and service controls while transparently engaging in HIE without a central intermediary. The output of this work, if implemented in a health system, will increase healthcare service quality and data quality through automated accountability. The system will ease regulatory functions like terminology database update distribution. Finally, Patient-centered care, now the hallmark of US Office of National Coordinator (ONC) strategy, becomes easy to implement because of the sharing economy and power facilitated by this work.

6.2 Academic significance

The evidence of state of art in software integration and Health Information Exchange (HIE) significantly contributes to knowledge. The development of a model for shared tokens in healthcare will prompt new research in healthcare and cross-sector business models leveraging data. This work also demystified the world of healthcare standards, particularly the Fast Healthcare Interoperability Resource (FHIR) standard.

6.3 Critique and Limitations

The literature search relied on scholarly works and expert community testimonials to determine that limited evidence of HIE implementation currently exists. It is possible that excluding non-scholarly, unpublished, or projects under different degrees of development may lead to the non-inclusion of specific unpublished initiatives. It is also important to clarify that in the context of this Thesis, sharing information within an organization is not considered a HIE. The potential impact of this is that when organizations share information within their enterprise, it may be construed to mean Health Information Exchange (HIE) and can mean trivializing the challenge involved in multi-stakeholder information exchange, thus limiting potential applicability of this research.

The health service providers' survey covered fewer facilities than expected. This may be interpreted as not representative, yet, the aim was to establish wether information is standardized, and the finding show that even for the few, information to be shared was not standardized. The SEERS cancer data used for simulating patient-data logging on a blockchain was not structured in FHIR format and will give better insight if restructured into the standard format. This will mean that the size could double when properly structured with other data elements not needed.

Also, the Reference implementation was tested on a computer on the same local network. In production, there are conditions needed for such high value information flow that may slightly impact the overall process, and as the resource utilization was measured, these additional conditions may not really impact its applicability as the simulation bandwidth is still small. Conditions such as authentication, authorizations, of health ecosystem participants, pseudonymization and anonymization of patient's Protected Health Information (PHI) can introduce complexities which can increase bandwidth utilization, data processing cost, and dependencies on external data sources, which can result in additional overheads. The study of these factors, which are beyond the scope of current research may have impact depending on the use case chosen. In addition, this work used the International Patient Summary (IPS) for the simulation. The IPS is a general purpose patient summary, for Patients with chronic illnesses or those that need specialized care, more information may be needed than currently captured in any IPS bundle. For instance, imaging use case requires a thousand times more storage and processing needs than other health use case encounters. This is a potential limitation of this proof of concept, yet, this research will recommend addressing Health Information Exchange (HIE) one use case at a time (eg. first hypertension, then immunization, and so on) for the greatest health impact. Other use cases may include additional stakeholder mapping, and resource considerations.

6.4 Future Work

Since this is the first shared health token in published literature, additional future work is needed to enhance practical implementation. In the future, the implementation of RegistryChain as a lightweight node and performance-tuned will greatly contribute to knowledge. Specifically, the academic community can extend this research through the following potential areas of future work:

- Standards adoption remain a challenge, for instance, even when two applications use the same version of FHIR, implementation often vary. Sometime, variation in the terminology (semantics) standards can also result further deviation from standards and eventual interoperability. Mandatory agreement within jurisdictions on implementation details for each use case is essential to appropriate HIE
- Adaptation of shared token model to one health use case, as against general purpose framework based on IPS as presented will enhance practical applicability
- It is also possible that imaging use case may best benefit from storing hashed image information on-chain and the actual image data stored off-chain for optimal applicability.
- Execution of additional tests for the tokens-based sharing economy will help stress test this novel healthcare token model and ready it for industrial application
- Implementation and simulation of these Token framework in a typical LMIC pilot program will help further support the hypothesis in practice
- Implementation of lightweight RegistryChain on a typical Internet of Things (IoT) hardware device like Raspberry Pi will further support the hypothesis of low resource consumption
- Experiment and test using different identity management models using different Certificate Authority (CA) architectures will definitively show the optimal CA for an LMIC
- Survey of health practitioners will be better validated (and considered more representative) if more health professionals are surveyed, possible across multiple jurisdictions
- Investigate how open-source software solutions compare to proprietary equivalents, especially in Low and Middle Income Countries (LMIC), contributing to the discussion on which is better, open source or proprietary
- Packaging and publishing an all-inclusive containerized version of RegistryChain will make it easier to install for low literate users

In this study, the different integration patterns were considered, and a comparison of document HIE and API-based exchange HIE will prove invaluable. These and more opportunities exist for expanding this work.

References

[1] E. Chukwu and L. Garg: "A systematic review of blockchain in healthcare: frameworks, prototypes, and implementations" *IEEE Access* pp. 1–1 ISSN 2169-3536 URL http://dx.doi.org/10.1109/ACCESS.2020.2969881, 2020

[2] E. Chukwu, L. Garg, E. Foday, A. Konomanyi, R. Wright, and F. Smart: "Digital health solutions and state of interoperability: Sierra Leone's Landscape analysis" *JMIR* URL http://dx.doi.org/10.2196/29930, 2022

[3] E. Chukwu, L. Garg, N. Obande-Ogbuinya, and V. K. Chattu: "Standardizing Primary Healthcare referral datasets: Interviews, form-reviews, and FHIR profiling" *JMIR formative research* URL http://dx.doi.org/10.2196/28510, 2022

[4] M. Dameron: "Beigepaper: An Ethereum Technical Specification" URL https://github.com/chronaeon/beigepaper/, 2019

[5] M. E. Conway: "How do committees invent" *Datamation* vol. 14(4) ISSN 00116963, 1968

[6] T. Benson and G. Grieve: "The Health Information Revolution" *in* "Principles of Health Interoperability: FHIR, HL7 and SNOMED CT", chap. 1, p. 6 Springer URL http://dx.doi.org/10.1007/978-3-319-30370-3_9, 2020

[7] P. Padmanabhan: "Is healthcare too hard for Big Tech firms?" URL https://www.healthcareitnews.com/blog/healthcare-too-hard-big-tech-firms, 2021

[8] G. Hohpe and B. Woolf: *Enterprise Integration Patterns* Addison-Wesley Professional, 2003

[9] NIBSS: "NIBSS" URL https://nibss-plc.com.ng, 2021

[10] J. M. Overhage: "Case Study 1 - The Indiana Health Information Exchange" *in* B. E. Dixon, editor, "Health Information Exchange", pp. 267–279 Academic Press ISBN 978-0-12-803135-3 URL http://dx.doi.org/https://doi.org/10.1016/B978-0-12-803135-3.00027-X, 2016

[11] T. Benson and G. Grieve: "SNOMED CT" *in* "Principles of Health Interoperability", Springer URL http://dx.doi.org/10.1007/978-3-319-30370-3_9, 2020

[12] International Health Terminology Standards Development Organization (IHTSDO): "Systematized Normenclature of Medicine – Clinical Terms (SNOMED-CT)" URL https://www.snomed.org, 2021

[13] Regenstrief Institute: "The International standard for identifying health measurements, observation, and documents (LOINC)" URL https://loinc.org, 2021

[14] World Health Organization: "International Classification of Diseases 11th Revision" URL https://icd.who.int/en/, 2021

[15] N. NLM: "RxNorm" *in* "Definitions", NIH NLM URL http://dx.doi.org/10.32388/a5ooex, 2020

[16] DICOM: "About DICOM: Overview" URL https://www.dicomstandard.org/about, 2020

[17] D. Capko, S. Vukmirovic, and N. Nedic: "State of the Art of Zero-Knowledge Proofs in Blockchain" *in* "2022 30th

Telecommunications Forum (TELFOR)", pp. 1–4 IEEE ISBN 978-1-6654-7273-9 URL http://dx.doi.org/10.1109/TELFOR56187.2022.9983760, 2022

[18] M. Fowler: "Richardson Maturity Model" URL https://martinfowler.com/articles/richardsonMaturityModel.html, 2010

[19] M. Fowler: *Patterns of Enterprise Application Architecture* vol. 48 Addison-Wesley Professional, 2002

[20] A. W. Brown: "Model driven architecture: Principles and practice" *Software and Systems Modeling* ISSN 1619-1366 URL http://dx.doi.org/10.1007/s10270-004-0061-2, 2004

[21] D. Emery and R. Hilliard: "Every architecture description needs a framework: Expressing architecture frameworks using ISO/IEC 42010" *in* "2009 Joint Working IEEE/IFIP Conference on Software Architecture and European Conference on Software Architecture, WICSA/ECSA 2009", ISBN 9781424449859 URL http://dx.doi.org/10.1109/WICSA.2009.5290789, 2009

[22] E. Gamma, R. Helm, R. Johnson, and J. Vlissides: *Design Patterns: Elements of Reusable Object-Oriented Software* Addison-Wesley Professional URL http://dx.doi.org/10.4324/9780203583159-18, 1994

[23] IEEE-SA Standards Board: "IEEE Recommended Practice for Architectural Description of Software-Intensive Systems" *IEEE Std*, 2000

[24] P. Bourque and R. E. Fairley: *SWEBOK v.3 - Guide to the Software Engineering - Body of Knowledge.* IEEE Computer Society ISBN 0-7695-2330-7 URL http://dx.doi.org/10.1234/12345678, 2014

[25] Object Management Group (OMG): "OMG Unified Modeling Language TM (OMG UML), Infrastructure 2.4.1" *InformatikSpektrum* ISSN 08950695 URL http://dx.doi.org/10.1007/s002870050092, 2011

[26] Object Management Group (OMG): "Business Process Model and Notation (BPMN) Version 2.0" *Business* ISSN 13507540 URL http://dx.doi.org/10.1007/s11576-008-0096-z, 2011

[27] SysML: "SysML Open Source Project - What is SysML? Who created SysML?" URL http://www.sysml.org, 2020

[28] M. Lim: "C2CFTP: Direct and Indirect File Transfer Protocols Between Clients in Client-Server Architecture" *IEEE Access* vol. 8:102833–102845 ISSN 2169-3536 URL http://dx.doi.org/10.1109/ACCESS.2020.2998725, 2020

[29] B. Bergsten, M. Couprie, and M. Lopez: "DBS3: a parallel database system for shared store" *in* "[1993] Proceedings of the Second International Conference on Parallel and Distributed Information Systems", pp. 260–262 IEEE Comput. Soc. Press ISBN 0-8186-3330-1 URL http://dx.doi.org/10.1109/PDIS.1993.253084, 1993

[30] Roy Thomas Fielding: "Fielding Dissertation: CHAPTER 5: Representational State Transfer (REST)" , 2000

[31] F. Liang, S. Liu, X. Meng, and C. Yang: "An Integrated Multi-channel Messaging Model supporting for business collaboration" *in* "Proceedings of the 2011 15th International Conference on Computer Supported Cooperative Work in Design (CSCWD)", pp. 532–537 IEEE ISBN 978-1-4577-0386-7 URL http://dx.doi.org/10.1109/CSCWD.2011.5960123, jun 2011

[32] S. Nakamoto: "Bitcoin: A Peer-to-Peer Electronic Cash" *Journal for General Philosophy of Science* vol. 39(1):53–67 ISSN 09254560 URL http://dx.doi.org/10.1007/s10838-008-9062-0, 2008

[33] V. Buterin: "Ethereum White Paper: A Next Generation Smart Contract & Decentralized Application Platform" URL http://dx.doi.org/10.5663/aps.v1i1.10138, 2013

[34] Hyperledger Performance and Scale Working Group: "Hyperledger Blockchain Performance Metrics" URL https://www.hyperledger.org/wp-content/uploads/2018/10/HL_Whitepaper_Metrics_PDF_V1.01.pdf, 2018

[35] S. Rahurkar, B. E. Dixon, and N. Menachemi: "Drivers and Barriers to Adoption: Towards the Last Mile" *in* B. E.

Dixon, editor, “Health Information Exchange: Navigating a Network of Health Information Systems”, chap. 3, p. 42 Academic Press ISBN 978-0-12-803135-3 URL http://dx.doi.org/10.1016/B978-0-12-803135-3.00003-7, 2016

[36] M. E. Kruk, A. D. Gage, N. T. Joseph, G. Danaei, S. García-Saisó, and J. A. Salomon: “Mortality due to low-quality health systems in the universal health coverage era: a systematic analysis of amenable deaths in 137 countries” *The Lancet* vol. 392(10160):2203–2212 ISSN 01406736 URL http://dx.doi.org/10.1016/S0140-6736(18)31668-4, nov 2018

[37] Grand View Research: “Healthcare IT Market Size, Share and Trends Analysis Report By Application (EHR, CPOE, Electronic prescribing systems, Medical Imaging Information), By Delivery Mode, By End Use, By Region, And Segment Forecasts, 2024 - 2030” URL https://www.grandviewresearch.com/industry-analysis/healthcare-it-market, 2023

[38] WHA: “Seventy-first World Health Assembly resolution A71.20/A/CONF./1 Agenda item: 12.4 Digital health” , 2018

[39] H. M. Scobie, M. Edelstein, E. Nicol, A. Morice, N. Rahimi, N. E. MacDonald, M. Carolina Danovaro-Holliday, and J. Jawad: “Improving the quality and use of immunization and surveillance data: Summary report of the Working Group of the Strategic Advisory Group of Experts on Immunization” *Vaccine* vol. 38(46):7183–7197 ISSN 0264410X URL http://dx.doi.org/10.1016/j.vaccine.2020.09.017, oct 2020

[40] W. R. Hersh, A. M. Totten, K. Eden, B. Devine, P. Gorman, S. Z. Kassakian, S. S. Woods, M. Daeges, M. Pappas, and M. S. McDonagh: “The Evidence Base for Health Information Exchange” *in* B. E. Dixon, editor, “Health Information Exchange”, chap. 3, pp. 213–229 Academic Press ISBN 978-0-12-803135-3 URL http://dx.doi.org/https://doi.org/10.1016/B978-0-12-803135-3.00014-1, 2016

[41] WHA: “Sixty-sixth World Health Assembly resolution A66.24/ Agenda item: eHealth standardization and Interoperability” , 2013

[42] Office of National Coordinator for Health Information Technology: “Connecting Health and Care of the Nation: A shared nationwide interoperability roadmap” Tech. rep. ONC HIT Washington DC URL https://www.healthit.gov/sites/default/files/nationwide-interoperability-roadmap-draft-version-1.0.pdf, 2019

[43] S. Ahern, R. Feiler, and S. Sdrinis: “Maximising the value of clinical registry information through integration with a health service clinical governance framework: a case study” *Australian Health Review* vol. 44:421 ISSN 0156-5788 URL http://dx.doi.org/10.1071/AH19004, 2020

[44] B. E. Dixon: “What is Health Information Exchange?” *in* B. E. Dixon, editor, “Health Information Exchange”, chap. 1, pp. 3–20 Academic Press ISBN 978-0-12-803135-3 URL http://dx.doi.org/https://doi.org/10.1016/B978-0-12-803135-3.00001-3, 2016

[45] M. Hosseini and B. E. Dixon: “Syntactic Interoperability and the Role of Standards” *in* B. E. Dixon, editor, “Health Information Exchange”, chap. 8, pp. 123–136 Academic Press ISBN 978-0-12-803135-3 URL http://dx.doi.org/10.1016/B978-0-12-803135-3.00008-6, 2016

[46] B. Séroussi and J. Bouaud: “The (Re)-Relaunching of the DMP, the French Shared Medical Record: New Features to Improve Uptake and Use.” *Studies in health technology and informatics* vol. 247:256–260 ISSN 1879-8365, 2018

[47] S. N. Haque, R. Bailey, and B. Massoudi: “Case Study 2 - Using Health Information Exchange to Support Public Health Activities in Western New York: A Case Study” *in* B. E. Dixon, editor, “Health Information Exchange”, pp. 281–294 Academic Press ISBN 978-0-12-803135-3 URL http://dx.doi.org/https://doi.org/10.1016/B978-0-12-803135-3.00028-1, 2016

[48] OpenHIE: “OpenHIE impact stories” URL https://ohie.org/impact-stories/, 2021

[49] K. W. Kelley, S. S. Feldman, and S. D. Gravely: “Engaging and Sustaining Stakeholders: Towards Gover-

nance" *in* B. Dixon, editor, "Health Information Exchange - Navigating and Managing a Network of Health Information Systems", 1st ed. chap. 4, pp. 59–76 Elsevier, Indianapolis URL http://dx.doi.org/10.1016/B978-0-12-803135-3.00004-9, 2016

[50] Y. Zhuang, L. R. Sheets, Y.-W. Chen, Z.-Y. Shae, J. J. Tsai, and C.-R. Shyu: "A Patient-Centric Health Information Exchange Framework Using Blockchain Technology" *IEEE Journal of Biomedical and Health Informatics* vol. 24(8):2169–2176 ISSN 2168-2194 URL http://dx.doi.org/10.1109/JBHI.2020.2993072, aug 2020

[51] A. P. Singh, N. R. Pradhan, A. K. Luhach, S. Agnihotri, N. Z. Jhanjhi, S. Verma, Kavita, U. Ghosh, and D. S. Roy: "A Novel Patient-Centric Architectural Framework for Blockchain-Enabled Healthcare Applications" *IEEE Transactions on Industrial Informatics* vol. 17:5779–5789 ISSN 1551-3203 URL http://dx.doi.org/10.1109/TII.2020.3037889, 2021

[52] Office of National Coordinator for Health Information Technology: "Connecting Health and Care of the Nation: A shared nationwide interoperability roadmap" Tech. rep. ONC Washington DC URL https://www.healthit.gov/sites/default/files/nationwide-interoperability-roadmap-draft-version-1.0.pdf, 2019

[53] HIMSS: "Interoperability in Healthcare" URL https://www.himss.org/resources/interoperability-healthcare, 2020

[54] T. J. Mowbray and R. Zahavi: *The essential CORBA : systems integration using distributed objects* na ISBN 0471106119, 1995

[55] J. M. Alyea, B. E. Dixon, J. Bowie, and A. S. Kanter: "Standardizing Health-Care Data Across an Enterprise" *in* B. E. Dixon, editor, "Health Information Exchange", chap. 9, pp. 137–148 Academic Press ISBN 978-0-12-803135-3 URL http://dx.doi.org/https://doi.org/10.1016/B978-0-12-803135-3.00009-8, 2016

[56] Health Level Seven: "HL7 EHR System Functional Model: A major development towards consensus on Electronic health Records System Functionality" URL https://www.hl7.org/documentcenter/public/wg/ehr/EHR-SWhitePaper.pdf, 2004

[57] HL7: "International Patient Summary Implementation Guide" URL http://hl7.org/fhir/uv/ips/, 2021

[58] Technical Committee ISO/TC 215: "The ISO 13606 standard" URL http://www.en13606.org/information.html, 2019

[59] OpenEHR: "Open EHR Architypes" URL https://www.openehr.org, 2019

[60] World Health Organization: "COVID-19 coding in ICD-10" Tech. rep. WHO, 2020

[61] E. Chukwu: "The role of digital ID in healthcare" *HealthTech law and regulation* pp. 167–192 URL http://dx.doi.org/10.4337/9781839104909.00018, 2020

[62] T. D. McFarlane, B. E. Dixon, and S. J. Grannis: "Client Registries: Identifying and Linking Patients" *in* B. E. Dixon, editor, "Health Information Exchange", chap. 11, pp. 163–182 Academic Press ISBN 978-0-12-803135-3 URL http://dx.doi.org/https://doi.org/10.1016/B978-0-12-803135-3.00011-6, 2016

[63] WHO: "Health Workforce" URL https://www.who.int/health-topics/health-workforce#tab=tab_1, 2023

[64] Federal Ministry of Health: "Nigeria Health Facility Registry (HFR)" URL https://hfr.health.gov.ng/statistics/tables, 2019

[65] 104th US congress: "Health Insurance Portability and Accountability Act (HIPAA)" , 1996

[66] Information Commissioner's Office: "Guide to the General Data Protection Regulation (GDPR)" *Guide to the General Data Protection Regulation* URL https://www.gov.uk/government/publications/guide-to-the-general-data-protection-regulation, 2018

[67] M. McNabb, E. Chukwu, H. Salami, O. Ojo, and F. Jega: "Assessing the feasibility and value of a pilot project using mobile applications and mobile money to enhance a maternal health conditional cash transfer (CCT) program in Nigeria leading to the development of a costed business model for scale up" *Annals of Global Health* vol. 81:193 ISSN 2214-9996 URL http://dx.doi.org/10.1016/j.aogh.2015.02.944, 2015

[68] Center for Disease Control and Prevention: "Meaningful Use of Electronic Health Records" URL https://www.cdc.gov/cancer/npcr/meaningful_use.htm, 2018

[69] A. Akhlaq, A. Sheikh, and C. Pagliari: "Health Information Exchange as a Complex and Adaptive Construct: Scoping Review" *Journal of Innovation in Health Informatics* vol. 23(4):633 ISSN 2058-4563 URL http://dx.doi.org/10.14236/jhi.v23i4.889, 2017

[70] I. Johansen, G. Henriksen, K. Demkjaer, H. B. Jensen, and L. Jørgensen: "Quality assurance and certification of health IT-systems communicating data in primary and secondary health sector." *Studies in health technology and informatics* vol. 95:601–5 ISSN 0926-9630, 2003

[71] N. ANGULA, N. DLODLO, and P. Q. MTSHALI: "Enabling Semantic Interoperability of Crowdsourced Disease Surveillance Data for Namibia Through a Health-Standards-Based Approach" *in* "2019 IST-Africa Week Conference (IST-Africa)", pp. 1–9 IEEE ISBN 978-1-905824-63-2 URL http://dx.doi.org/10.23919/ISTAFRICA.2019.8764830, 2019

[72] J. L. Gor: "Agent-based interoperability system in health insurance" *in* "2017 IST-Africa Week Conference (IST-Africa)", pp. 1–9 IEEE URL http://dx.doi.org/10.23919/ISTAFRICA.2017.8101978, 2017

[73] F. Khalique, S. A. Khan, and I. Nosheen: "A Framework for Public Health Monitoring, Analytics and Research" *IEEE Access* vol. 7:101309–101326 ISSN 2169-3536 URL http://dx.doi.org/10.1109/ACCESS.2019.2930730, 2019

[74] A. Akhlaq, B. McKinstry, and A. Sheikh: "Stakeholders perspectives and deployment strategies of health information exchange illustrated through an in-depth case study of Pakistan" *Informatics for Health and Social Care* vol. 45(2):130–150 ISSN 1753-8157 URL http://dx.doi.org/10.1080/17538157.2019.1582053, 2020

[75] C. Seebregts, P. Dane, A. N. Parsons, T. Fogwill, D. Rogers, M. Bekker, V. Shaw, and P. Barron: "Designing for scale: optimising the health information system architecture for mobile maternal health messaging in South Africa (MomConnect)" *BMJ Global Health* vol. 3:e000563 ISSN 2059-7908 URL http://dx.doi.org/10.1136/bmjgh-2017-000563, 2018

[76] IntraHealth International: "iHRIS" , 2021

[77] WHO European Region: "E-health in practice" URL https://www.euro.who.int/en/countries/estonia/news/news/016/03/e-health-in-practice, 2016

[78] eHealth Network: "Exchange of electronic health records across the EU" URL https://digital-strategy.ec.europa.eu/en/policies/electronic-health-records, 2021

[79] InterSystems: "Netherlands Health Information Exchange Wins Trust of Millions" Tech. rep. Intersystems, 2019

[80] A. J. Holmgren and J. Adler-Milstein: "Health Information Exchange in U.S. Hospitals: The Current Landscape and a Path to Improved Information Sharing" *Journal of Hospital Medicine* vol. 12(03):193–198 ISSN 15535606 URL http://dx.doi.org/10.12788/jhm.2704, 2017

[81] Healthix: "Healthix" URL https://healthix.org, 2021

[82] IHIE: "Indiana Health Information Exhange (IHIE)" URL https://www.ihie.org, 2020

[83] HealthEConnections: "Health Information Exchange (HIE) Services" URL https://www.healtheconnections.org/what-we-do/hie-services/, 2021

[84] na: "Midwest health connection" URL https://www.mhc-hie.org, 2021

[85] Healthie Nevada: "Healthie Nevada" URL https://healthienevada.org/participants/, 2021

[86] W. G. van Panhuis, P. Paul, C. Emerson, J. Grefenstette, R. Wilder, A. J. Herbst, D. Heymann, and D. S. Burke: "A systematic review of barriers to data sharing in public health" *BMC Public Health* vol. 14(1):1144 ISSN 1471-2458 URL http://dx.doi.org/10.1186/1471-2458-14-1144, 2014

[87] R. Chandrasekaran, B. Sankaranarayanan, and J. Pendergrass: "Unfulfilled promises of health information exchange: What inhibits ambulatory clinics from electronically sharing health information?" *International Journal of Medical Informatics* vol. 149:104418 ISSN 13865056 URL http://dx.doi.org/10.1016/j.ijmedinf.2021.104418, 2021

[88] C. De Pietro and I. Francetic: "E-health in Switzerland: The laborious adoption of the federal law on electronic health records (EHR) and health information exchange (HIE) networks" *Health Policy* vol. 122(2):69–74 ISSN 01688510 URL http://dx.doi.org/10.1016/j.healthpol.2017.11.005, 2018

[89] M. M. MELLO, J. ADLER-MILSTEIN, K. L. DING, and L. SAVAGE: "Legal Barriers to the Growth of Health Information Exchange-Boulders or Pebbles?" *The Milbank Quarterly* vol. 96(1):110–143 ISSN 0887378X URL http://dx.doi.org/10.1111/1468-0009.12313, 2018

[90] H. Ji, S. Yoo, E.-Y. Heo, H. Hwang, and J.-W. Kim: "Technology and Policy Challenges in the Adoption and Operation of Health Information Exchange Systems" *Healthcare Informatics Research* vol. 23(4):314 ISSN 2093-3681 URL http://dx.doi.org/10.4258/hir.2017.23.4.314, 2017

[91] W. R. Hersh, A. M. Totten, K. Eden, B. Devine, P. Gorman, S. Z. Kassakian, S. S. Woods, M. Daeges, M. Pappas, and M. S. McDonagh: "The Evidence Base for Health Information Exchange" *in* B. E. Dixon, editor, "Health Information Exchange", chap. 14, pp. 213–229 Academic Press ISBN 978-0-12-803135-3 URL http://dx.doi.org/https://doi.org/10.1016/B978-0-12-803135-3.00014-1, 2016

[92] M. J. Dobrow, J. P. Bytautas, S. Tharmalingam, and S. Hagens: "Interoperable Electronic Health Records and Health Information Exchanges: Systematic Review" *JMIR Medical Informatics* vol. 7(2):e12607 ISSN 2291-9694 URL http://dx.doi.org/10.2196/12607, 2019

[93] S. E. Ross, T. A. Radcliff, W. G. LeBlanc, L. M. Dickinson, A. M. Libby, and D. E. Nease: "Effects of health information exchange adoption on ambulatory testing rates" *Journal of the American Medical Informatics Association* vol. 20:1137–1142 ISSN 1067-5027 URL http://dx.doi.org/10.1136/amiajnl-2012-001608, 2013

[94] C. M. Carr, C. S. Gilman, D. M. Krywko, H. E. Moore, B. J. Walker, and S. H. Saef: "Observational Study and Estimate of Cost Savings from Use of a Health Information Exchange in an Academic Emergency Department" *The Journal of Emergency Medicine* vol. 46:250–256 ISSN 07364679 URL http://dx.doi.org/10.1016/j.jemermed.2013.05.068, 2014

[95] M. E. Frisse, K. B. Johnson, H. Nian, C. L. Davison, C. S. Gadd, K. M. Unertl, P. A. Turri, and Q. Chen: "The financial impact of health information exchange on emergency department care" *Journal of the American Medical Informatics Association* vol. 19(3):328–333 ISSN 1067-5027 URL http://dx.doi.org/10.1136/amiajnl-2011-000394, 2012

[96] A. Tzeel, V. Lawnicki, and K. R. Pemble: "The business case for payer support of a community-based health information exchange: a human pilot evaluating its effectiveness in cost control for plan members seeking emergency department care." *American health and drug benefits* vol. 4:207–16 ISSN 1942-2962, 2011

[97] D. M. Walker: "Does participation in health information exchange improve hospital efficiency?" *Health Care Management Science* vol. 21(3):426–438 ISSN 1386-9620 URL http://dx.doi.org/10.1007/s10729-017-9396-4, 2018

[98] Technical Committee ISO/TC 215: "ISO 12052:2017(en), Health informatics — Digital imaging and communication in medicine (DICOM) including workflow and data management" , 2017

[99] Health Level Seven: "HL7-Home" URL http://www.hl7.org.uk/, 2020

[100] CEN: "CEN - EN 1064" *in* "Health informatics - Standard communication protocol - Computer-assisted electrocardiography", CEN URL https://standards.globalspec.com/std/14322011/en-1064, 2020

[101] CEN: "BSI - BS DD ENV 13607" *in* "Health Informatics - Messages for the Exchange of Information on Medicine Prescriptions", CEN URL https://standards.globalspec.com/std/599074/bs-dd-env-13607, 2020

[102] ISO: "ISO/CD TS 22691" *in* "Health informatics — Token-based health information sharing", ISO URL https://www.iso.org/standard/73692.html, 2021

[103] ISO: "ISO/DIS 27799" *in* "Health informatics— Information security management in health using ISO/IEC 27002", ISO, 2016

[104] CEN: "CEN/TR 15299:2006" *in* "TC 251 - Health informatics - Safety procedures for identification of patients and related objects", CEN, 2006

[105] CEN: "EN 14485:2003" *in* "Health informatics - Guidance for handling personal health data in international applications in the context of the EU data protection directive", CEN, 2017

[106] K. Osei-Tutu and Y.-T. Song: "Enterprise Architecture for Healthcare Information Exchange (HIE) Cloud Migration" *in* "2020 14th International Conference on Ubiquitous Information Management and Communication (IMCOM)", pp. 1–8 IEEE ISBN 978-1-7281-5453-4 URL http://dx.doi.org/10.1109/IMCOM48794.2020.9001677, 2020

[107] E. Felizmenio and S. Festin: "Evaluating the Viability of Ciphertext-Policy Attribute-Based Encryption on Service-Oriented Health Information Exchange Systems" *in* "2019 Eleventh International Conference on Ubiquitous and Future Networks (ICUFN)", pp. 443–448 IEEE ISBN 978-1-7281-1340-1 URL http://dx.doi.org/10.1109/ICUFN.2019.8806133, 2019

[108] J. Angala, A. C. Dacoco, E. Felizmenio, and W. M. Tan: "Leveraging Software-Defined Networking (SDN) Capabilities For Improving Health Information Exchange (HIE) Systems Performance" *in* "TENCON 2018 - 2018 IEEE Region 10 Conference", pp. 1743–1748 IEEE ISBN 978-1-5386-5457-6 URL http://dx.doi.org/10.1109/TENCON.2018.8650132, oct 2018

[109] Y. Yang, X. Li, N. Qamar, P. Liu, W. Ke, B. Shen, and Z. Liu: "Medshare: A Novel Hybrid Cloud for Medical Resource Sharing Among Autonomous Healthcare Providers" *IEEE Access* vol. 6:46949–46961 ISSN 2169-3536 URL http://dx.doi.org/10.1109/ACCESS.2018.2865535, 2018

[110] M. A. Uddin, A. Stranier, I. Gondal, and V. Balasubramanian: "A patient agent to manage blockchains for remote patient monitoring" *in* "Studies in Health Technology and Informatics", ISBN 9781614999133 ISSN 18798365 URL http://dx.doi.org/10.3233/978-1-61499-914-0-105, 2018

[111] J. P. Kansky: "Managing the Business of Health Information Exchange: Toward Sustainability" *in* B. E. Dixon, editor, "Health Information Exchange", chap. 5, pp. 77–89 Academic Press ISBN 978-0-12-803135-3 URL http://dx.doi.org/https://doi.org/10.1016/B978-0-12-803135-3.00005-0, 2016

[112] Nordic Institute for Interoperability Solutions: "X-ROAD" URL https://x-road.global, 2021

[113] X. Zhou, V. Jesus, Y. Wang, and M. Josephs: "User-Controlled, Auditable, Cross-Jurisdiction Sharing of Healthcare Data Mediated by a Public Blockchain" *in* "2020 IEEE 19th International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom)", pp. 87–96 IEEE ISBN 978-1-6654-0392-4 URL http://dx.doi.org/10.1109/TrustCom50675.2020.00025, dec 2020

[114] S. Amofa, E. B. Sifah, K. O. . O. Agyekum, S. Abla, Q. Xia, J. C. Gee, and J. Gao: "A Blockchain-based Architecture Framework for Secure Sharing of Personal Health Data" *in* "2018 IEEE 20th International Conference on e-Health Networking, Applications and Services (Healthcom)", pp. 1–6 URL http://dx.doi.org/10.1109/HealthCom.2018.8531160, 2018

[115] N. A. Azeez and C. V. der Vyver: "Dynamic Patient-Regulated Access Control Framework for Electronic Health Information" *in* "2017 International Conference on Computational Science and Computational Intelligence (CSCI)",

pp. 1684–1690 IEEE ISBN 978-1-5386-2652-8 URL http://dx.doi.org/10.1109/CSCI.2017.293, dec 2017

[116] J. Vest and T. Miller: "The association between health information exchange and measures of patient satisfaction" *Applied Clinical Informatics* vol. 02:447–459 ISSN 1869-0327 URL http://dx.doi.org/10.4338/ACI-2011-06-RA-0040, 2011

[117] Y. Zhuang, L. R. Sheets, Y.-W. Chen, Z.-Y. Shae, J. J. Tsai, and C.-R. Shyu: "A Patient-Centric Health Information Exchange Framework Using Blockchain Technology" *IEEE Journal of Biomedical and Health Informatics* vol. 24(8):2169–2176 URL http://dx.doi.org/10.1109/JBHI.2020.2993072, 2020

[118] A. Roehrs, C. A. da Costa, and R. da Rosa Righi: "OmniPHR: A distributed architecture model to integrate personal health records" *Journal of Biomedical Informatics* ISSN 15320464 URL http://dx.doi.org/10.1016/j.jbi.2017.05.012, 2017

[119] C. Burniske, E. Vaughn, J. Shelton, and A. Cahana: "How Blockchain Technology Can Enhance EHR Operability" *Gem | Ark Invest Research*, 2016

[120] A. Azaria, A. Ekblaw, T. Vieira, and A. Lippman: "MedRec: Using blockchain for medical data access and permission management" *in* "Proceedings - 2016 2nd International Conference on Open and Big Data, OBD 2016", ISBN 9781509040544 ISSN 9781509040544 URL http://dx.doi.org/10.1109/OBD.2016.11, 2016

[121] D. Ichikawa, M. Kashiyama, and T. Ueno: "Tamper-Resistant Mobile Health Using Blockchain Technology" *JMIR mHealth and uHealth* ISSN 2291-5222 URL http://dx.doi.org/10.2196/mhealth.7938, 2017

[122] L. Castaldo and V. Cinque: "Blockchain-based logging for the cross-border exchange of ehealth data in Europe" *in* "Communications in Computer and Information Science", ISBN 9783319951881 ISSN 18650929 URL http://dx.doi.org/10.1007/978-3-319-95189-8_5, 2018

[123] K. N. Griggs, O. Ossipova, C. P. Kohlios, A. N. Baccarini, E. A. Howson, and T. Hayajneh: "Healthcare Blockchain System Using Smart Contracts for Secure Automated Remote Patient Monitoring" *Journal of Medical Systems* ISSN 1573689X URL http://dx.doi.org/10.1007/s10916-018-0982-x, 2018

[124] A. Dubovitskaya, Z. Xu, S. Ryu, M. Schumacher, and F. Wang: "How blockchain could empower ehealth: An application for radiation oncology: (Extended abstract)" *in* "Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)", ISBN 9783319671857 ISSN 16113349 URL http://dx.doi.org/10.1007/978-3-319-67186-4_1, 2017

[125] A. Ekblaw and Azaria Asaf: "MedRec: Medical Data Management on the Blockchain · PubPub" URL http://dx.doi.org/10.1103/PhysRevA.60.4693, 2016

[126] A. Allison, C. Anne, N. Diakun-Thibault, L. Forni, F. Landa, J. Mayo, and R. Vab Riezen: "Blockchain and Health it: Algorithms, Privacy and Data" Tech. rep. MIT, 2016

[127] J. H. Tseng, Y. C. Liao, B. Chong, and S. W. Liao: "Governance on the drug supply chain via gcoin blockchain" *International Journal of Environmental Research and Public Health* ISSN 16604601 URL http://dx.doi.org/10.3390/ijerph15061055, 2018

[128] S. Jiang, J. Cao, H. Wu, Y. Yang, M. Ma, and J. He: "Blochie: A blockchain-based platform for healthcare information exchange" *in* "Proceedings - 2018 IEEE International Conference on Smart Computing, SMARTCOMP 2018", ISBN 9781538647059 URL http://dx.doi.org/10.1109/SMARTCOMP.2018.00073, 2018

[129] Y. Ji, J. Zhang, J. Ma, C. Yang, and X. Yao: "BMPLS: Blockchain-Based Multi-level Privacy-Preserving Location Sharing Scheme for Telecare Medical Information Systems" *Journal of Medical Systems* ISSN 1573689X URL http://dx.doi.org/10.1007/s10916-018-0998-2, 2018

[130] K. Fan, S. Wang, Y. Ren, H. Li, and Y. Yang: "MedBlock: Efficient and Secure Medical Data Sharing Via Blockchain" *Journal of Medical Systems* ISSN 1573689X URL http://dx.doi.org/10.1007/s10916-018-0993-7, 2018

[131] X. Yue, H. Wang, D. Jin, M. Li, and W. Jiang: "Healthcare Data Gateways: Found Healthcare Intelligence on Blockchain with Novel Privacy Risk Control" *Journal of Medical Systems* ISSN 1573689X URL http://dx.doi.org/10.1007/s10916-016-0574-6, 2016

[132] Y. Sun, R. Zhang, X. Wang, K. Gao, and L. Liu: "A Decentralizing Attribute-Based Signature for Healthcare Blockchain" *in* "2018 27th International Conference on Computer Communication and Networks (ICCCN)", ISBN 978-1-5386-5156-8 ISSN 1095-2055 URL http://dx.doi.org/10.1109/ICCCN.2018.8487349, 2018

[133] M. A. Rahman, M. S. Hossain, G. Loukas, E. Hassanain, S. S. Rahman, M. F. Alhamid, and M. Guizani: "Blockchain-Based Mobile Edge Computing Framework for Secure Therapy Applications" *IEEE Access* vol. 6:72469–72478 ISSN 2169-3536 VO - 6 URL http://dx.doi.org/10.1109/ACCESS.2018.2881246, 2018

[134] J. Liu, X. Li, L. Ye, H. Zhang, X. Du, and M. Guizani: "BPDS: A Blockchain based Privacy-Preserving Data Sharing for Electronic Medical Records" *CoRR* vol. abs/1811.0, 2018

[135] G. Nagasubramanian, R. K. Sakthivel, R. Patan, A. H. Gandomi, and M. Sankayya: "Securing e-health records using keyless signature infrastructure blockchain technology in the cloud" *Neural Computing and Applications* pp. 1–9, 2018

[136] A. Ekblaw, A. Azaria, J. D. Halamka, and A. Lippman: "A Case Study for Blockchain in Healthcare: "MedRec"prototype for electronic health records and medical research data" *Proceedings of IEEE Open & Big Data Conference* ISSN 0888-7543 URL http://dx.doi.org/10.1006/geno.1994.1313, 2016

[137] Q. Xia, E. B. Sifah, K. O. Asamoah, J. Gao, X. Du, and M. Guizani: "MeDShare: Trust-Less Medical Data Sharing among Cloud Service Providers via Blockchain" *IEEE Access* ISSN 21693536 URL http://dx.doi.org/10.1109/ACCESS.2017.2730843, 2017

[138] Q. Xia, E. B. Sifah, A. Smahi, S. Amofa, and X. Zhang: "BBDS: Blockchain-based data sharing for electronic medical records in cloud environments" *Information (Switzerland)* ISSN 20782489 URL http://dx.doi.org/10.3390/info8020044, 2017

[139] A. D. MS, Z. Xu, S. Ryu, and M. Schumacher: "Secure and Trustable Electronic Medical Records Sharing using Blockchain." *2017 AMIA Annual Symposium proceedings*, 2018

[140] T. Bocek, B. B. Rodrigues, T. Strasser, and B. Stiller: "Blockchains everywhere - A use-case of blockchains in the pharma supply-chain" *in* "Proceedings of the IM 2017 - 2017 IFIP/IEEE International Symposium on Integrated Network and Service Management", ISBN 9783901882890 ISSN 00098981 URL http://dx.doi.org/10.23919/INM.2017.7987376, 2017

[141] M. Benchoufi, R. Porcher, and P. Ravaud: "Blockchain protocols in clinical trials: Transparency and traceability of consent" *F1000Research* ISSN 2046-1402 URL http://dx.doi.org/10.12688/f1000research.10531.4, 2017

[142] S.-J. Lee, G.-Y. Cho, F. Ikeno, and T.-R. Lee: "BAQALC: Blockchain Applied Lossless Efficient Transmission of DNA Sequencing Data for Next Generation Medical Informatics" *Applied Sciences* ISSN 2076-3417 URL http://dx.doi.org/10.3390/app8091471, 2018

[143] A. Roehrs, C. A. da Costa, R. da Rosa Righi, V. F. da Silva, J. R. Goldim, and D. C. Schmidt: "Analyzing the performance of a blockchain-based personal health record implementation" *Journal of Biomedical Informatics* vol. 92:103140 ISSN 15320464 URL http://dx.doi.org/10.1016/j.jbi.2019.103140, 2019

[144] R. C. Celiz, Y. E. De La Cruz, and D. M. Sanchez: "Cloud Model for Purchase Management in Health Sector of Peru based on IoT and Blockchain" *in* "2018 IEEE 9th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON)", pp. 328–334 IEEE ISBN 978-1-5386-7266-2 URL http://dx.doi.org/10.1109/IEMCON.2018.8615063, 2018

[145] J. H. Brenas, M. S. Al-Manir, C. J. O. Baker, and A. Shaban-Nejad: "A Malaria Analytics Framework to Support Evolution and Interoperability of Global Health Surveillance Systems" *IEEE Access* vol. 5:21605–21619 ISSN 2169-

3536 URL http://dx.doi.org/10.1109/ACCESS.2017.2761232, 2017

[146] ISO: "ISO/TS 21089:2018" *in* "Health informatics — Trusted end-to-end information flows", ISO URL https://www.iso.org/standard/66936.html, 2018

[147] ISO: "ISO/IEC TR 14516:2002" *in* "Information technology — Security techniques — Guidelines for the use and management of Trusted Third Party services", ISO URL https://www.iso.org/standard/31482.html, 2002

[148] ISO: "ISO 17090-1:2013" *in* "Health informatics — Public key infrastructure — Part 1: Overview of digital certificate services", ISO URL https://www.iso.org/standard/63019.html, 2013

[149] ISO: "ISO/IEC 10118-3:2018" *in* "IT Security techniques — Hash-functions — Part 3: Dedicated hash-functions", ISO URL https://www.iso.org/standard/67116.html, 2018

[150] CEN: "CR 13694:1999" *in* "CEN/TC- 251 - Health Informatics - Safety and Security Related Software Quality Standards for Healthcare (SSQS)", CEN, 2000

[151] ISO: "ISO/TR 21730:2007" *in* "Health informatics — Use of mobile wireless communication and computing technology in healthcare facilities — Recommendations for electromagnetic compatibility (management of unintentional electromagnetic interference) with medical devices", ISO URL https://www.iso.org/standard/44865.html, 2007

[152] C. técnico ISO: "ISO 18308:2011 Health Informatics – Requirements for an Electronic Health Record Architecture" *Health (San Francisco)*, 2011

[153] A. Roehrs, C. Andr, R. Righi, V. Ferreira, R. Goldim, and D. C. Schmidt: "Analyzing the Performance of a Blockchain-based Personal Health Record Implementation" *Journal of Biomedical Informatics* vol. 00(00):1–10 URL http://dx.doi.org/10.1016/j.jbi.2019.103140, 2018

[154] M. Hölbl, M. Kompara, and A. Kamišali: "A Systematic Review of the Use of Blockchain in Healthcare" *MDPI* pp. 1–22 ISSN 2073-8994 URL http://dx.doi.org/10.20944/preprints201809.0136.v1, 2018

[155] K. Peterson, R. Deeduvanu, P. Kanjamala, and K. Boles: "A blockchain-based approach to health information exchange networks" *Proc. NIST Workshop Blockchain Healthcare* ISSN 18770509 URL http://dx.doi.org/10.1016/j.procs.2015.08.363, 2016

[156] T. Nugent, D. Upton, and M. Cimpoesu: "Improving data transparency in clinical trials using blockchain smart contracts" *F1000Research* ISSN 2046-1402 URL http://dx.doi.org/10.12688/f1000research.9756.1, 2016

[157] G. Magyar: "Blockchain: Solving the privacy and research availability tradeoff for EHR data: A new disruptive technology in health data management" *in* "IEEE 30th Jubilee Neumann Colloquium, NC 2017", ISBN 9781538646366 ISSN 0930-7575 URL http://dx.doi.org/10.1109/NC.2017.8263269, 2018

[158] J. Brogan, I. Baskaran, and N. Ramachandran: "Authenticating Health Activity Data Using Distributed Ledger Technologies" *Computational and Structural Biotechnology Journal* vol. 16:257–266 ISSN 20010370 URL http://dx.doi.org/10.1016/j.csbj.2018.06.004, 2018

[159] S. H. Lee and C. S. Yang: "Fingernail analysis management system using microscopy sensor and blockchain technology" *International Journal of Distributed Sensor Networks* ISSN 15501477 URL http://dx.doi.org/10.1177/1550147718767044, 2018

[160] F. Angeletti, I. Chatzigiannakis, and A. Vitaletti: "The role of blockchain and IoT in recruiting participants for digital clinical trials" *in* "2017 25th International Conference on Software, Telecommunications and Computer Networks, SoftCOM 2017", ISBN 9789532900781 ISSN 0031-4005 URL http://dx.doi.org/10.23919/SOFTCOM.2017.8115590, 2017

[161] A. Zhang and X. Lin: "Towards Secure and Privacy-Preserving Data Sharing in e-Health Systems via Consortium Blockchain" *Journal of Medical Systems* ISSN 1573689X URL http://dx.doi.org/10.1007/

s10916-018-0995-5, 2018

[162] P. Sylim, F. Liu, A. Marcelo, and P. Fontelo: "Blockchain technology for detecting falsified and substandard drugs in distribution: Pharmaceutical supply chain intervention" *Journal of Medical Internet Research* ISSN 14388871 URL http://dx.doi.org/10.2196/10163, 2018

[163] T. Dey, S. Jaiswal, S. Sunderkrishnan, and N. Katre: "HealthSense: A medical use case of Internet of Things and blockchain" *in* "Proceedings of the International Conference on Intelligent Sustainable Systems, ICISS 2017", ISBN 9781538619599 URL http://dx.doi.org/10.1109/ISS1.2017.8389459, 2018

[164] H. Wang and Y. Song: "Secure Cloud-Based EHR System Using Attribute-Based Cryptosystem and Blockchain" *Journal of Medical Systems* ISSN 1573689X URL http://dx.doi.org/10.1007/s10916-018-0994-6, 2018

[165] X. Liang, S. Shetty, J. Zhao, D. Bowden, D. Li, and J. Liu: "Towards decentralized accountability and self-sovereignty in healthcare systems" *in* "Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)", ISBN 9783319894997 ISSN 16113349 URL http://dx.doi.org/10.1007/978-3-319-89500-0_34, 2018

[166] X. Zheng, R. R. Mukkamala, R. Vatrapu, and J. Ordieres-Mere: "Blockchain-based Personal Health Data Sharing System Using Cloud Storage" *in* "2018 IEEE 20th International Conference on e-Health Networking, Applications and Services (Healthcom)", pp. 1–6 URL http://dx.doi.org/10.1109/HealthCom.2018.8531125, 2018

[167] X. Liang, J. Zhao, S. Shetty, J. Liu, and D. Li: "Integrating blockchain for data sharing and collaboration in mobile healthcare applications" *in* "IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, PIMRC", ISBN 9781538635315 ISSN 01406736 URL http://dx.doi.org/10.1109/PIMRC.2017.8292361, 2018

[168] M. A. Uddin, A. Stranieri, I. Gondal, and V. Balasubramanian: "Continuous Patient Monitoring With a Patient Centric Agent: A Block Architecture" *IEEE Access* ISSN 2169-3536 URL http://dx.doi.org/10.1109/ACCESS.2018.2846779, 2018

[169] L. Zhou, L. Wang, and Y. Sun: "MIStore: a Blockchain-Based Medical Insurance Storage System" *Journal of Medical Systems* ISSN 1573689X URL http://dx.doi.org/10.1007/s10916-018-0996-4, 2018

[170] J. Chen, X. Ma, M. Du, and Z. Wang: "A Blockchain Application for Medical Information Sharing" *in* "TEMS-ISIE 2018 - 1st Annual International Symposium on Innovation and Entrepreneurship of the IEEE Technology and Engineering Management Society", ISBN 9781538644751 URL http://dx.doi.org/10.1109/TEMS-ISIE.2018.8478645, 2018

[171] K. Mannaro, G. Baralla, A. Pinna, and S. Ibba: "A blockchain approach applied to a teledermatology platform in the Sardinian Region (Italy)" *Information (Switzerland)* ISSN 20782489 URL http://dx.doi.org/10.3390/info9020044, 2018

[172] H. Jiang, H. Peng, and S. Dian: "A design of medical information sharing model based on blockchain technology" *IOP Conference Series: Materials Science and Engineering*, 2018

[173] Y. Du, J. Liu, Z. Guan, and H. Feng: "A Medical Information Service Platform Based on Distributed Cloud and Blockchain" *in* "2018 IEEE International Conference on Smart Cloud (SmartCloud)", pp. 34–39 URL http://dx.doi.org/10.1109/SmartCloud.2018.00014, 2018

[174] A. C., K. M., and N. Mohananthini: "A secured healthcare system using private blockchain technology" *Journal of Engineering technology* vol. 6(2):42–54, 2018

[175] G. Srivastava, A. D. Dwivedi, and R. Singh: "Automated Remote Patient Monitoring: Data Sharing and Privacy Using Blockchain." *arXiv Computers and Society* URL http://dx.doi.org/arXiv:1811.03417v1, 2018

[176] S. Rahmadika and K. H. Rhee: "Blockchain technology for providing an architecture model of decentralized personal health information" *International Journal of Engineering Business Management* ISSN 18479790 URL http://dx.doi.org/10.1177/1847979018790589, 2018

[177] M. Z. A. Bhuiyan, A. Zaman, T. Wang, G. Wang, H. Tao, and M. M. Hassan: "Blockchain and Big Data to Transform the Healthcare" *in* "Proceedings of the International Conference on Data Processing and Applications", ICDPA 2018 pp. 62–68 ACM, New York, NY, USA ISBN 978-1-4503-6418-8 URL http://dx.doi.org/10.1145/3224207.3224220, 2018

[178] T. T. Thwin and S. Vasupongayya: "Blockchain Based Secret-Data Sharing Model for Personal Health Record System" *in* "2018 5th International Conference on Advanced Informatics: Concept Theory and Applications (ICAICTA)", pp. 196–201 URL http://dx.doi.org/10.1109/ICAICTA.2018.8541296, 2018

[179] S. Alexaki, G. Alexandris, V. Katos, and E. N. Petroulakis: "Blockchain-based Electronic Patient Records for Regulated Circular Healthcare Jurisdictions" *in* "2018 IEEE 23rd International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD)", pp. 1–6 URL http://dx.doi.org/10.1109/CAMAD.2018.8514954, 2018

[180] J. Liu: "HIV digital vaccine strategy : An application of blockchain technology in preventing the spread of HIV" *pre proposal*, 2018

[181] S. L. Cichosz, M. N. Stausholm, T. Kronborg, P. Vestergaard, and O. Hejlesen: "How to Use Blockchain for Diabetes Health Care Data and Access Management: An Operational Concept" URL http://dx.doi.org/10.1177/1932296818790281, 2018

[182] M. Hanley and H. Tewari: "Managing Lifetime Healthcare Data on the Blockchain" *in* "2018 IEEE SmartWorld, Ubiquitous Intelligence and Computing, Advanced and Trusted Computing, Scalable Computing and Communications, Cloud and Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI)", pp. 246–251 URL http://dx.doi.org/10.1109/SmartWorld.2018.00077, 2018

[183] P. B. Nichol and W. Dailey: "Micro-Identities Improve Healthcare Interoperability with Blockchain: Deterministic Methods for Connecting Patient Data to Uniform Patient Identifiers" *ResearchGate*, 2016

[184] A. Theodouli, S. Arakliotis, K. Moschou, K. Votis, and D. Tzovaras: "On the Design of a Blockchain-Based System to Facilitate Healthcare Data Sharing" *in* "Proceedings - 17th IEEE International Conference on Trust, Security and Privacy in Computing and Communications and 12th IEEE International Conference on Big Data Science and Engineering, Trustcom/BigDataSE 2018", ISBN 9781538643877 ISSN 2324-9013 URL http://dx.doi.org/10.1109/TrustCom/BigDataSE.2018.00190, 2018

[185] A. K. Talukder, M. Chaitanya, D. Arnold, and K. Sakurai: "Proof of Disease: A Blockchain Consensus Protocol for Accurate Medical Decisions and Reducing the Disease Burden" *in* "2018 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI)", pp. 257–262 ISBN VO - URL http://dx.doi.org/10.1109/SmartWorld.2018.00079, 2018

[186] Z. Shae and J. J. Tsai: "On the Design of a Blockchain Platform for Clinical Trial and Precision Medicine" *in* "Proceedings - International Conference on Distributed Computing Systems", ISBN 9781538617915 ISSN 1063-6927 URL http://dx.doi.org/10.1109/ICDCS.2017.61, 2017

[187] A. Al Omar, M. S. Rahman, A. Basu, and S. Kiyomoto: "MediBchain: A blockchain based privacy preserving platform for healthcare data" *in* "Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)", ISBN 9783319723945 ISSN 16113349 URL http://dx.doi.org/10.1007/978-3-319-72395-2_49, 2017

[188] T.-T. Kuo and L. Ohno-Machado: "ModelChain: Decentralized Privacy-Preserving Healthcare Predictive Modeling Framework on Private Blockchain Networks" *CoRR* vol. abs/1802.0, 2018

[189] G. G. Dagher, J. Mohler, M. Milojkovic, and P. B. Marella: "Ancile: Privacy-preserving framework for access control and interoperability of electronic health records using blockchain technology" *Sustainable Cities and Society* ISSN 22106707 URL http://dx.doi.org/10.1016/j.scs.2018.02.014, 2018

[190] P. Zhang, J. White, D. C. Schmidt, G. Lenz, and S. T. Rosenbloom: "FHIRChain: Applying Blockchain to Securely and Scalably Share Clinical Data" *Computational and Structural Biotechnology Journal* ISSN 20010370 URL http://dx.doi.org/10.1016/j.csbj.2018.07.004, 2018

[191] W. Liu, S. S. Zhu, T. Mundie, and U. Krieger: "Advanced block-chain architecture for e-health systems" *in* "2017 IEEE 19th International Conference on e-Health Networking, Applications and Services, Healthcom 2017", ISBN 9781509067046 URL http://dx.doi.org/10.1109/HealthCom.2017.8210847, 2017

[192] P. Mangesius, J. Bachmann, T. Healy, S. Saboor, and T. Schabetsberger: "Blockchains in IHE-Based Networks." *Studies in health technology and informatics* ISSN 1879-8365, 2018

[193] S. Badr, I. Gomaa, and E. Abd-Elrahman: "Multi-tier Blockchain Framework for IoT-EHRs Systems" *Procedia Computer Science* ISSN 18770509 URL http://dx.doi.org/10.1016/j.procs.2018.10.162, 2018

[194] A. F. Hussein, N. ArunKumar, G. Ramirez-Gonzalez, E. Abdulhay, J. M. R. Tavares, and V. H. C. de Albuquerque: "A medical records managing and securing blockchain based system supported by a Genetic Algorithm and Discrete Wavelet Transform" *Cognitive Systems Research* ISSN 13890417 URL http://dx.doi.org/10.1016/j.cogsys.2018.05.004, 2018

[195] R. Guo, H. Shi, Q. Zhao, and D. Zheng: "Secure Attribute-Based Signature Scheme with Multiple Authorities for Blockchain in Electronic Health Records Systems" *IEEE Access* ISSN 21693536 URL http://dx.doi.org/10.1109/ACCESS.2018.2801266, 2018

[196] J. P. Dias, L. Reis, H. S. Ferreira, and A. Martins: "Blockchain for Access Control in e-Health Scenarios" *CoRR* vol. abs/1805.1, 2018

[197] M. L. Gagnon and G. Stephen: "A Pragmatic Solution to a Major Interoperability Problem: Using Blockchain for the Nationwide Patient Index" Tech. rep. na URL https://blockchainhealthcaretoday.com/index.php/journal/article/view/28, 2018

[198] H. Yang and B. Yang: "A Blockchain-based Approach to the Secure Sharing of Healthcare Data" *in* "Norwgian Information Security Conference", , 2017

[199] V. Patel: "A framework for secure and decentralized sharing of medical imaging data via blockchain consensus" URL http://dx.doi.org/10.1177/1460458218769699, 2018

[200] H.-J. Cha, H.-K. Yang, and Y.-J. Song: "A Study on Access Structure Management of CP-ABTD Based Blockchain for Medical Information Monitoring System" *Advanced Science Letters* ISSN 1936-6612 URL http://dx.doi.org/10.1166/asl.2018.11838, 2018

[201] E. Gökalp, M. O. Gökalp, S. Çoban, and P. E. Eren: "Analysing opportunities and challenges of integrated blockchain technologies in healthcare" *in* "Lecture Notes in Business Information Processing", ISBN 9783030000592 ISSN 18651348 URL http://dx.doi.org/10.1007/978-3-030-00060-8_13, 2018

[202] E. Funk, J. Riddell, F. Ankel, and D. Cabrera: "Blockchain Technology: A Data Framework to Improve Validity, Trust, and Accountability of Information Exchange in Health Professions Education" *Academic medicine : journal of the Association of American Medical Colleges* ISSN 1938808X URL http://dx.doi.org/10.1097/ACM.0000000000002326, 2018

[203] Y. Chen, S. Ding, Z. Xu, H. Zheng, and S. Yang: "Blockchain-Based Medical Records Secure Storage and Medical Service Framework" *Journal of Medical Systems* ISSN 1573689X URL http://dx.doi.org/10.1007/s10916-018-1121-4, 2018

[204] Noh S.: "Blockchain-based user-centric records management system" *International Journal of Control and Automation*, 2017

[205] M. A. Rahman, M. S. Hossain, M. M. Rashid, S. J. Barnes, M. F. Alhamid, and M. Guizani: "A Blockchain-Based Non-Invasive Cyber-Physical Occupational Therapy Framework: BCI Perspective" *IEEE Access* vol. 7:34874–34884 ISSN 2169-3536 URL http://dx.doi.org/10.1109/ACCESS.2019.2903024, 2019

[206] Z. Gao, L. Xu, G. Turner, B. Patel, N. Diallo, L. Chen, and W. Shi: "Blockchain-based Identity Management with Mobile Device" *in* "Proceedings of the 1st Workshop on Cryptocurrencies and Blockchains for Distributed Systems - CryBlock'18", pp. 66–70 ACM Press, New York, New York, USA ISBN 9781450358385 URL http://dx.doi.org/10.1145/3211933.3211945, 2018

[207] S. Chakraborty, S. Aich, and H.-C. Kim: "A Secure Healthcare System Design Framework using Blockchain Technology" *in* "2019 21st International Conference on Advanced Communication Technology (ICACT)", pp. 260–264 IEEE ISBN 979-11-88428-02-1 URL http://dx.doi.org/10.23919/ICACT.2019.8701983, feb 2019

[208] K. Ito, K. Tago, and Q. Jin: "i-Blockchain: A Blockchain-Empowered Individual-Centric Framework for Privacy-Preserved Use of Personal Health Data" *in* "2018 9th International Conference on Information Technology in Medicine and Education (ITME)", pp. 829–833 IEEE ISBN 978-1-5386-7744-5 URL http://dx.doi.org/10.1109/ITME.2018.00186, oct 2018

[209] J. Vora, A. Nayyar, S. Tanwar, S. Tyagi, N. Kumar, M. S. Obaidat, and J. J. P. C. Rodrigues: "BHEEM: A Blockchain-Based Framework for Securing Electronic Health Records" *in* "2018 IEEE Globecom Workshops (GC Wkshps)", pp. 1–6 IEEE ISBN 978-1-5386-4920-6 URL http://dx.doi.org/10.1109/GLOCOMW.2018.8644088, dec 2018

[210] V. Ramani, T. Kumar, A. Bracken, M. Liyanage, and M. Ylianttila: "Secure and Efficient Data Accessibility in Blockchain Based Healthcare Systems" *in* "2018 IEEE Global Communications Conference (GLOBECOM)", pp. 206–212 IEEE ISBN 978-1-5386-4727-1 URL http://dx.doi.org/10.1109/GLOCOM.2018.8647221, dec 2018

[211] R. Wutthikarn and Y. G. Hui: "Prototype of Blockchain in Dental care service application based on Hyperledger Composer in Hyperledger Fabric framework" *in* "2018 22nd International Computer Science and Engineering Conference (ICSEC)", pp. 1–4 IEEE ISBN 978-1-5386-8164-0 URL http://dx.doi.org/10.1109/ICSEC.2018.8712639, nov 2018

[212] E. Zaghloul, T. Li, and J. Ren: "Security and Privacy of Electronic Health Records: Decentralized and Hierarchical Data Sharing using Smart Contracts" *in* "2019 International Conference on Computing, Networking and Communications (ICNC)", pp. 375–379 IEEE ISBN 978-1-5386-9223-3 URL http://dx.doi.org/10.1109/ICCNC.2019.8685552, feb 2019

[213] A. S. Kleinaki, P. Mytis-Gkometh, G. Drosatos, P. S. Efraimidis, and E. Kaldoudi: "A Blockchain-Based Notarization Service for Biomedical Knowledge Retrieval" *Computational and Structural Biotechnology Journal* ISSN 20010370 URL http://dx.doi.org/10.1016/j.csbj.2018.08.002, 2018

[214] H. Kaur, M. A. Alam, R. Jameel, A. K. Mourya, and V. Chang: "A Proposed Solution and Future Direction for Blockchain-Based Heterogeneous Medicare Data in Cloud Environment" *Journal of Medical Systems* ISSN 1573689X URL http://dx.doi.org/10.1007/s10916-018-1007-5, 2018

[215] H. Han, M. Huang, Y. Zhang, and U. A. Bhatti: "An architecture of secure health information storage system based on blockchain technology" *in* "Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)", ISBN 9783030000080 ISSN 16113349 URL http://dx.doi.org/10.1007/978-3-030-00009-7_52, 2018

[216] Z. Shae and J. J. Tsai: "Transform blockchain into distributed parallel computing architecture for precision medicine" *in* "Proceedings - International Conference on Distributed Computing Systems", ISBN 9781538668719 ISSN 1558-1950 URL http://dx.doi.org/10.1109/ICDCS.2018.00129, 2018

[217] L. Chen, W.-K. Lee, C.-C. Chang, K.-K. R. Choo, and N. Zhang: "Blockchain based searchable encryption for electronic health record sharing" *Future Generation Computer Systems* vol. 95:420–429 ISSN 0167739X URL http://dx.doi.org/10.1016/j.future.2019.01.018, 2019

[218] D. R. Ortega, C. M. Oikonomou, H. Jane Ding, P. Rees-Lee, Alexandria, and G. J. Jensen: "ETDB-Caltech: A blockchain-based distributed public database for electron tomography" *PLoS ONE* ISSN 19326203 URL http://dx.doi.org/10.1371/journal.pone.0215531, 2019

[219] K. Azbeg, O. Ouchetto, S. J. Andaloussi, L. Fetjah, and A. Sekkaki: "Blockchain and IoT for Security and Privacy: A Platform for Diabetes Self-management" *in* "2018 4th International Conference on Cloud Computing Technologies and Applications (Cloudtech)", pp. 1–5 IEEE ISBN 978-1-7281-1637-2 URL http://dx.doi.org/10.1109/CloudTech.2018.8713343, nov 2018

[220] S. Kim and D. Kim: "Design of an innovative blood cold chain management system using blockchain technologies" *ICIC Express Letters, Part B: Applications* ISSN 21852766 URL http://dx.doi.org/10.24507/icicelb.09.10.1067, 2018

[221] R. Kumar and R. Tripathi: "Traceability of counterfeit medicine supply chain through Blockchain" *in* "2019 11th International Conference on Communication Systems and Networks (COMSNETS)", pp. 568–570 IEEE ISBN 978-1-5386-7902-9 URL http://dx.doi.org/10.1109/COMSNETS.2019.8711418, jan 2019

[222] B. M. Till, A. W. Peters, S. Afshar, and J. Meara: "From blockchain technology to global health equity: can cryptocurrencies finance universal health coverage?" *BMJ Global Health* ISSN 2059-7908 URL http://dx.doi.org/10.1136/bmjgh-2017-000570, 2017

[223] T. Heston: "A case study in blockchain health care innovation" *International Journal of Current Research* vol. 9(11):60587 – 60588, 2017

[224] H. Tang, N. Tong, and J. Ouyang: "Medical Images Sharing System Based on Blockchain and Smart Contract of Credit Scores" *in* "2018 1st IEEE International Conference on Hot Information-Centric Networking (HotICN)", pp. 240–241 IEEE ISBN 978-1-5386-4870-4 URL http://dx.doi.org/10.1109/HOTICN.2018.8605956, aug 2018

[225] S. Zhangy, A. Kim, D. Liu, S. C. Nuckchadyy, L. Huangy, A. Masurkary, J. Zhangy, L. P. Karnatiz, L. Martinezx, T. Hardjono, M. Kellis, and Z. Zhang: "Genie: A Secure, Transparent Sharing and Services Platform for Genetic and Health Data" *eprint*, 2018

[226] K. Fan, S. Wang, Y. Ren, H. Li, and Y. Yang: "MedBlock: Efficient and Secure Medical Data Sharing Via Blockchain" *Journal of Medical Systems* ISSN 1573689X URL http://dx.doi.org/10.1007/s10916-018-0993-7, 2018

[227] A. Juneja and M. Marefat: "Leveraging blockchain for retraining deep learning architecture in patient-specific arrhythmia classification" *in* "2018 IEEE EMBS International Conference on Biomedical and Health Informatics, BHI 2018", ISBN 9781538624050 URL http://dx.doi.org/10.1109/BHI.2018.8333451, 2018

[228] X. Zhang and S. Poslad: "Blockchain Support for Flexible Queries with Granular Access Control to Electronic Medical Records (EMR)" *in* "2018 IEEE International Conference on Communications (ICC)", pp. 1–6 ISBN 1938-1883 VO URL http://dx.doi.org/10.1109/ICC.2018.8422883, 2018

[229] T. Mikula and R. H. Jacobsen: "Identity and access management with blockchain in electronic healthcare records" *in* "Proceedings - 21st Euromicro Conference on Digital System Design, DSD 2018", ISBN 9781538673768 URL http://dx.doi.org/10.1109/DSD.2018.00008, 2018

[230] I. Kotsiuba, A. Velvkzhanin, Y. Yanovich, I. S. Bandurova, Y. Dyachenko, and V. Zhygulin: "Decentralized e-Health Architecture for Boosting Healthcare Analytics" *in* "2018 Second World Conference on Smart Trends in Systems, Security and Sustainability (WorldS4)", pp. 113–118 IEEE ISBN 978-1-5386-7280-8 URL http://dx.doi.org/10.1109/WorldS4.2018.8611621, oct 2018

[231] X. Zhang, S. Poslad, and Z. Ma: "Block-Based Access Control for Blockchain-Based Electronic Medical Records (EMRs) Query in eHealth" *in* "2018 IEEE Global Communications Conference (GLOBECOM)", pp. 1–7 IEEE ISBN 978-1-5386-4727-1 URL http://dx.doi.org/10.1109/GLOCOM.2018.8647433, 2018

[232] M. Franceschi, D. Morelli, D. Plans, A. Brown, J. Collomosse, L. Coutts, and L. Ricci: "ComeHere: Exploiting Ethereum for Secure Sharing of Health-Care Data" *European Conference on Parallel Processing* pp. 585–596 URL http://dx.doi.org/10.1007/978-3-030-10549-5_46, 2019

[233] H. Tian, J. He, and Y. Ding: "Medical Data Management on Blockchain with Privacy" *Journal of Medical Systems*

vol. 43(2):26 ISSN 0148-5598 URL http://dx.doi.org/10.1007/s10916-018-1144-x, 2019

[234] S. Cao, G. Zhang, P. Liu, X. Zhang, and F. Neri: "Cloud-assisted secure eHealth systems for tamper-proofing EHR via blockchain" *Information Sciences* vol. 485:427–440 ISSN 00200255 URL http://dx.doi.org/10.1016/j.ins.2019.02.038, jun 2019

[235] R. N. Nortey, L. Yue, P. R. Agdedanu, and M. Adjeisah: "Privacy Module for Distributed Electronic Health Records(EHRs) Using the Blockchain" *in* "2019 IEEE 4th International Conference on Big Data Analytics (ICBDA)", pp. 369–374 IEEE ISBN 978-1-7281-1282-4 URL http://dx.doi.org/10.1109/ICBDA.2019.8713188, mar 2019

[236] A. R. Rajput, Q. Li, M. T. Ahvanooey, and I. Masood: "EACMS: Emergency Access Control Management System for Personal Health Record based on Blockchain" *IEEE Access* pp. 1–1 ISSN 2169-3536 URL http://dx.doi.org/10.1109/ACCESS.2019.2917976, 2019

[237] T. H.-J. Kim and J. Lampkins: "BRICS: Blockchain-based Resilient Information Control System" *in* "2018 IEEE International Conference on Big Data (Big Data)", pp. 5363–5365 IEEE ISBN 978-1-5386-5035-6 URL http://dx.doi.org/10.1109/BigData.2018.8621993, 2018

[238] T. Rupasinghe, F. Burstein, C. Rudolph, and S. Strange: "Towards a Blockchain based Fall Prediction Model for Aged Care" *in* "Proceedings of the Australasian Computer Science Week Multiconference", pp. 1–10 ACM, New York, NY, USA ISBN 9781450366038 URL http://dx.doi.org/10.1145/3290688.3290736, jan 2019

[239] A. Iyengar, A. Kundu, U. Sharma, and P. Zhang: "A trusted healthcare data analytics cloud platform" *in* "Proceedings - International Conference on Distributed Computing Systems", ISBN 9781538668719 URL http://dx.doi.org/10.1109/ICDCS.2018.00123, 2018

[240] S. Wang, J. Wang, X. Wang, T. Qiu, Y. Yuan, L. Ouyang, Y. Guo, and F. Y. Wang: "Blockchain-Powered Parallel Healthcare Systems Based on the ACP Approach" URL http://dx.doi.org/10.1109/TCSS.2018.2865526, 2018

[241] H. Li, L. Zhu, M. Shen, F. Gao, X. Tao, and S. Liu: "Blockchain-Based Data Preservation System for Medical Data" *Journal of Medical Systems* ISSN 1573689X URL http://dx.doi.org/10.1007/s10916-018-0997-3, 2018

[242] P. Zhang, J. White, D. C. Schmidt, and G. Lenz: "Applying Software Patterns to Address Interoperability Challenges in Blockchain-based Healthcare Apps" Yes, 2017

[243] Q. Xia, E. B. Sifah, K. O. Asamoah, J. Gao, X. Du, and M. Guizani: "MeDShare: Trust-Less Medical Data Sharing among Cloud Service Providers via Blockchain" *IEEE Access* ISSN 21693536 URL http://dx.doi.org/10.1109/ACCESS.2017.2730843, 2017

[244] H. L. Pham, T. H. Tran, and Y. Nakashima: "A Secure Remote Healthcare System for Hospital Using Blockchain Smart Contract" *in* "2018 IEEE Globecom Workshops (GC Wkshps)", pp. 1–6 IEEE ISBN 978-1-5386-4920-6 URL http://dx.doi.org/10.1109/GLOCOMW.2018.8644164, 2018

[245] A. Mohsin, A. Zaidan, B. Zaidan, O. Albahri, A. Albahri, M. Alsalem, and K. Mohammed: "Based blockchain-PSO-AES techniques in finger vein biometrics: A novel verification secure framework for patient authentication" *Computer Standards & Interfaces* ISSN 09205489 URL http://dx.doi.org/10.1016/j.csi.2019.04.002, 2019

[246] M. A. Rahman, E. Hassanain, M. M. Rashid, S. J. Barnes, and M. S. Hossain: "Spatial Blockchain-Based Secure Mass Screening Framework for Children With Dyslexia" *IEEE Access* vol. 6:61876–61885 ISSN 2169-3536 VO - 6 URL http://dx.doi.org/10.1109/ACCESS.2018.2875242, 2018

[247] T. Bocek, B. B. Rodrigues, T. Strasser, and B. Stiller: "Blockchains everywhere - A use-case of blockchains in the pharma supply-chain" *in* "Proceedings of the IM 2017 - 2017 IFIP/IEEE International Symposium on Integrated Network and Service Management", ISBN 9783901882890 ISSN 00098981 URL http://dx.doi.org/10.23919/INM.2017.7987376, 2017

[248] X. Li, P. Jiang, T. Chen, X. Luo, and Q. Wen: "A survey on the security of blockchain systems" URL http://dx.

doi.org/10.1016/j.future.2017.08.020, 2017

[249] Y. Ji, J. Zhang, J. Ma, C. Yang, and X. Yao: "BMPLS: Blockchain-Based Multi-level Privacy-Preserving Location Sharing Scheme for Telecare Medical Information Systems" *Journal of Medical Systems* ISSN 1573689X URL http://dx.doi.org/10.1007/s10916-018-0998-2, 2018

[250] T. Heston: "A Case Study in Blockchain Healthcare Innovation" *International Journal of Current Research* URL http://dx.doi.org/10.22541/au.151060471.10755953, 2017

[251] e-Governance Academy Tallinn: "e-Estonia e-Governance in Practice" Tech. rep. e-Governance Academy Tallinn Tallinn, 2019

[252] A. Novek: "Blockchain in Estonian National Health Information System" *Presentation* URL https://cdn.ymaws.com/echalliance.com/resource/resmgr/images/DHW_2018_Profiles_/2018_Presentations_/Artur_Novek.pdf, 2018

[253] Hyperledger Performance and Scale Working Group: "Hyperledger Blockchain Performance Metrics" URL https://www.hyperledger.org/wp-content/uploads/2018/10/HL_Whitepaper_Metrics_PDF_V1.01.pdf working group report, 2018

[254] H. Li, L. Zhu, M. Shen, F. Gao, X. Tao, and S. Liu: "Blockchain-Based Data Preservation System for Medical Data" *Journal of Medical Systems* ISSN 1573689X URL http://dx.doi.org/10.1007/s10916-018-0997-3, 2018

[255] X. Feng, J. Ma, T. Feng, Y. Miao, and X. Liu: "Consortium Blockchain-Based SIFT: Outsourcing Encrypted Feature Extraction in the D2D Network" *IEEE Access* vol. 6:52248–52260 ISSN 2169-3536 VO - 6 URL http://dx.doi.org/10.1109/ACCESS.2018.2869856, 2018

[256] M. A. Rahman, E. Hassanain, M. M. Rashid, S. J. Barnes, and M. S. Hossain: "Spatial Blockchain-Based Secure Mass Screening Framework for Children With Dyslexia" *IEEE Access* vol. 6:61876–61885 ISSN 2169-3536 VO - 6 URL http://dx.doi.org/10.1109/ACCESS.2018.2875242, 2018

[257] F. Tang, S. Ma, Y. Xiang, and C. Lin: "An Efficient Authentication Scheme for Blockchain-Based Electronic Health Records" *IEEE Access* vol. 7:41678–41689 ISSN 2169-3536 URL http://dx.doi.org/10.1109/ACCESS.2019.2904300, 2019

[258] Z. Xiao, Z. Li, Y. Liu, L. Feng, W. Zhang, T. Lertwuthikarn, and R. S. Mong Goh: "EMRShare: A Cross-Organizational Medical Data Sharing and Management Framework Using Permissioned Blockchain" *in* "2018 IEEE 24th International Conference on Parallel and Distributed Systems (ICPADS)", pp. 998–1003 IEEE ISBN 978-1-5386-7308-9 URL http://dx.doi.org/10.1109/PADSW.2018.8645049, 2018

[259] D. C. Nguyen, P. N. Pathirana, M. Ding, and A. Seneviratne: "Blockchain for Secure EHRs Sharing of Mobile Cloud Based E-Health Systems" *IEEE Access* vol. 7:66792–66806 ISSN 2169-3536 URL http://dx.doi.org/10.1109/ACCESS.2019.2917555, 2019

[260] M. Alblooshi, K. Salah, and Y. Alhammadi: "Blockchain-based Ownership Management for Medical IoT (MIoT) Devices" *in* "2018 International Conference on Innovations in Information Technology (IIT)", pp. 151–156 IEEE ISBN 978-1-5386-6673-9 URL http://dx.doi.org/10.1109/INNOVATIONS.2018.8606032, 2018

[261] T. Quaini, A. Roehrs, C. A. Da Costa, and R. Da Rosa Righi: "UNiReC: An architecture proposal for integrating distributed electronic health records using blockchain" *in* "Proceedings of the International Conferences on WWW/Internet 2018 and Applied Computing 2018", , 2018

[262] R. Lamba, Y. Gupta, S. Kalra, and M. Sharma: "Preventing Waiting List Manipulation And Black Marketing of Donated Organs Through Hyperledger Fabric" *in* "2019 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS)", pp. 280–285 IEEE ISBN 978-1-7281-4826-7 URL http://dx.doi.org/10.1109/ICCCIS48478.2019.8974526, 2019

[263] X. He, S. Alqahtani, and R. Gamble: "Toward Privacy-Assured Health Insurance Claims" *in* "2018 IEEE International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE

Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData)", pp. 1634–1641 IEEE ISBN 978-1-5386-7975-3 URL http://dx.doi.org/10.1109/Cybermatics_2018.2018.00273, jul 2018

[264] R. Raj, N. Rai, and S. Agarwal: "Anticounterfeiting in Pharmaceutical Supply Chain by establishing Proof of Ownership" *in* "TENCON 2019 - 2019 IEEE Region 10 Conference (TENCON)", pp. 1572–1577 IEEE ISBN 978-1-7281-1895-6 URL http://dx.doi.org/10.1109/TENCON.2019.8929271, oct 2019

[265] T. D. McFarlane, B. E. Dixon, and S. J. Grannis: "Client Registries: Identifying and Linking Patients" *in* B. Dixon, editor, "HEALTH INFORMATION EXCHANGE: navigating and managing a network of health information systems", 1st ed. chap. 11, pp. 163 – 182 Elsevier Inc. ISBN 978-0-12-803135-3, 2016

[266] e-Governance Academy Tallinn: "e-Estonia e-Governance in Practice" Tech. rep. e-Governance Academy Tallinn Tallinn, 2019

[267] Dalberg: "The State of Aahdaar: A People's Perspective" Tech. rep. Dalberg URL https://stateofaadhaar.in/assets/download/SoA_2019_Report_web.pdf, 2019

[268] Na: "PRISMA statement" URL http://www.prisma-statement.org, 2015

[269] L. Garg, E. Chukwu, N. Nasser, C. Chakraborty, and G. Garg: "Anonymity preserving IoT-based COVID-19 and other infectious disease contact tracing model" *IEEE Access* URL http://dx.doi.org/10.1109/ACCESS.2020.3020513, 2020

[270] Calculator.net: "Sample Size Calculator" URL https://www.calculator.net/sample-size-calculator.html, 2021

[271] M. de Bayser and R. Cerqueira: "Integrating MPI with Docker for HPC" *in* "2017 IEEE International Conference on Cloud Engineering (IC2E)", pp. 259–265 IEEE ISBN 978-1-5090-5817-4 URL http://dx.doi.org/10.1109/IC2E.2017.40, apr 2017

[272] WHO: *WORLD HEALTH STATISTICS 2018: Monitoring health for the SDGs* WHO, Geneva ISBN 9789241565585 URL http://apps.who.int/iris/bitstream/handle/10665/272596/9789241565585-eng.pdf?ua=1, 2018

[273] DPPI - MoHS: "National digital health strategy (2018 - 2023)" URL https://mohs2017.files.wordpress.com/2019/02/sl-national-digital-health-strategy-nov-2018.pdf, 2018

[274] F. Vogelsteller and V. Buterin: "Ethereum Improvement proposal" URL https://eips.ethereum.org/EIPS/eip-20, 2015

[275] S. R. P. r. A. . b. o. t. N. . s. National Cancer Institute DCCPS: "SEER*Stat Database: Incidence - SEER Research Data 8 Registries Nov 2021 Sub (1975-2019)" , 2023

[276] E. V. Bernstam, J. L. Warner, J. C. Krauss, E. Ambinder, W. S. Rubinstein, G. Komatsoulis, R. S. Miller, and J. L. Chen: "Quantitating and assessing interoperability between electronic health records" *Journal of the American Medical Informatics Association* vol. 29(5):753–760 URL http://dx.doi.org/10.1093/jamia/ocab289, 2022

[277] WHO: "World Health Statistics 2018: Monitoring the SDGs" Tech. rep. World Health Organization, 2018

[278] A. A. Bhattacharya, N. Umar, A. Audu, H. Felix, E. Allen, J. R. M. Schellenberg, and T. Marchant: "Quality of routine facility data for monitoring priority maternal and newborn indicators in DHIS2: A case study from Gombe State, Nigeria" *PLOS ONE* vol. 14(1):e0211265 ISSN 1932-6203 URL http://dx.doi.org/10.1371/journal.pone.0211265, 2019

[279] E. Chukwu, L. Garg, E. Foday, A. Konomanyi, R. Wright, and F. Smart: "Sierra Leone's health facilities' electricity, computing-hardware, and internet infrastructures: Field mapping" *JMIR Medical Informatics* ISSN 2291-9694 URL http://dx.doi.org/10.2196/30040, 2021

[280] M. Ali, A. Abbas, M. U. S. Khan, and S. U. Khan: "SeSPHR: A Methodology for Secure Sharing of Personal Health

Records in the Cloud" *IEEE Transactions on Cloud Computing* pp. 347–359 ISSN 2168-7161 URL http://dx.doi.org/10.1109/TCC.2018.2854790, 2021

[281] S. Osebe, A. Walcott, K. Weldemariam, C. M. Wachira, F. Matu, N. Bore, D. Kaguma, J. Mutahi, W. Ogallo, C. Cintas, and S. L. Remy: "Enabling Care Continuity using a Digital Health Wallet" *in* "2019 IEEE International Conference on Healthcare Informatics (ICHI)", pp. 1–7 IEEE ISBN 978-1-5386-9138-0 URL http://dx.doi.org/10.1109/ICHI.2019.8904625, jun 2019

[282] H. Singhal, H. Ravi, S. N. Chakravarthy, P. Balasundaram, and C. Babu: "EPMS: A Framework for Large-Scale Patient Matching" *in* "2019 IEEE 31st International Conference on Tools with Artificial Intelligence (ICTAI)", pp. 1096–1101 IEEE ISBN 978-1-7281-3798-8 URL http://dx.doi.org/10.1109/ICTAI.2019.00153, nov 2019

[283] Z. Ma, M. Zhao, X. Liu, C. Shen, and J. Ma: "Cross-Organizational Access Control for EHRs: Trustworthy, Flexible, Transparent" *in* "2019 IEEE Global Communications Conference (GLOBECOM)", pp. 1–6 IEEE ISBN 978-1-7281-0962-6 URL http://dx.doi.org/10.1109/GLOBECOM38437.2019.9013842, dec 2019

A Appendix - setup

Here, the steps to setup the reference model using Hyperledger Fabric is discussed. First, the data models section is used to discuss implementation of the Fast Healthcare Interoperability Resource (FHIR) reference SNOMED CT change management management schemas. The network section then discuss setup of RegistryChain blockchain orderers, peers, and smart contracts using Hyperledger fabric blockchain framework. The decision points and key configuration options for the Hyperledger network were detailed. An evaluation of the system is then discussed with key results.

A.1 Data models

In this section, the setup and implementation of a FHIR and SNOMED CT oracle, DApps, and smart contracts are described.

A.1.1 Off-chain FHIR oracle

Oracles in the blockchain context provide trusted services to blockchain data or service consumers. The data structure for money in FHIR is represented as in Figure A.1.

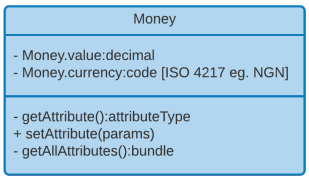


Figure A.1: Representation of in FHIR demonstrated for use with tokens

A.1.2 Usecase diagram

Based on the entire research, the, a usecase diagram was drawn to guide the framework development.

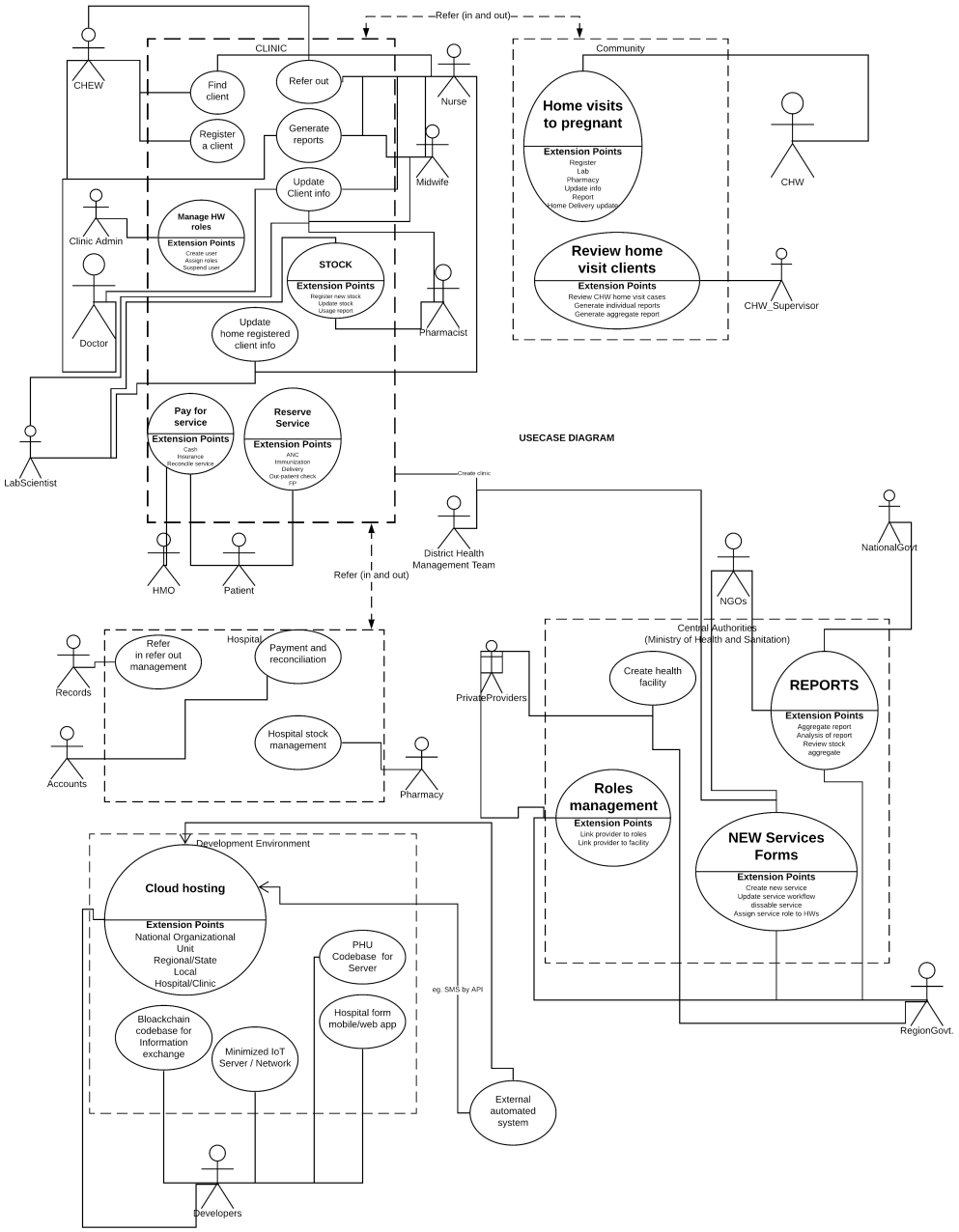


Figure A.2: A multi-stakeholder digital health use case (self-drawn)

The Ontology for the off-chain component of RegistryChain framework has been shown

in Figure A.3. Similarly, for the Organization, Practitioner and PractitionerRole FHIR syntactic report is shown in Figure A.4.

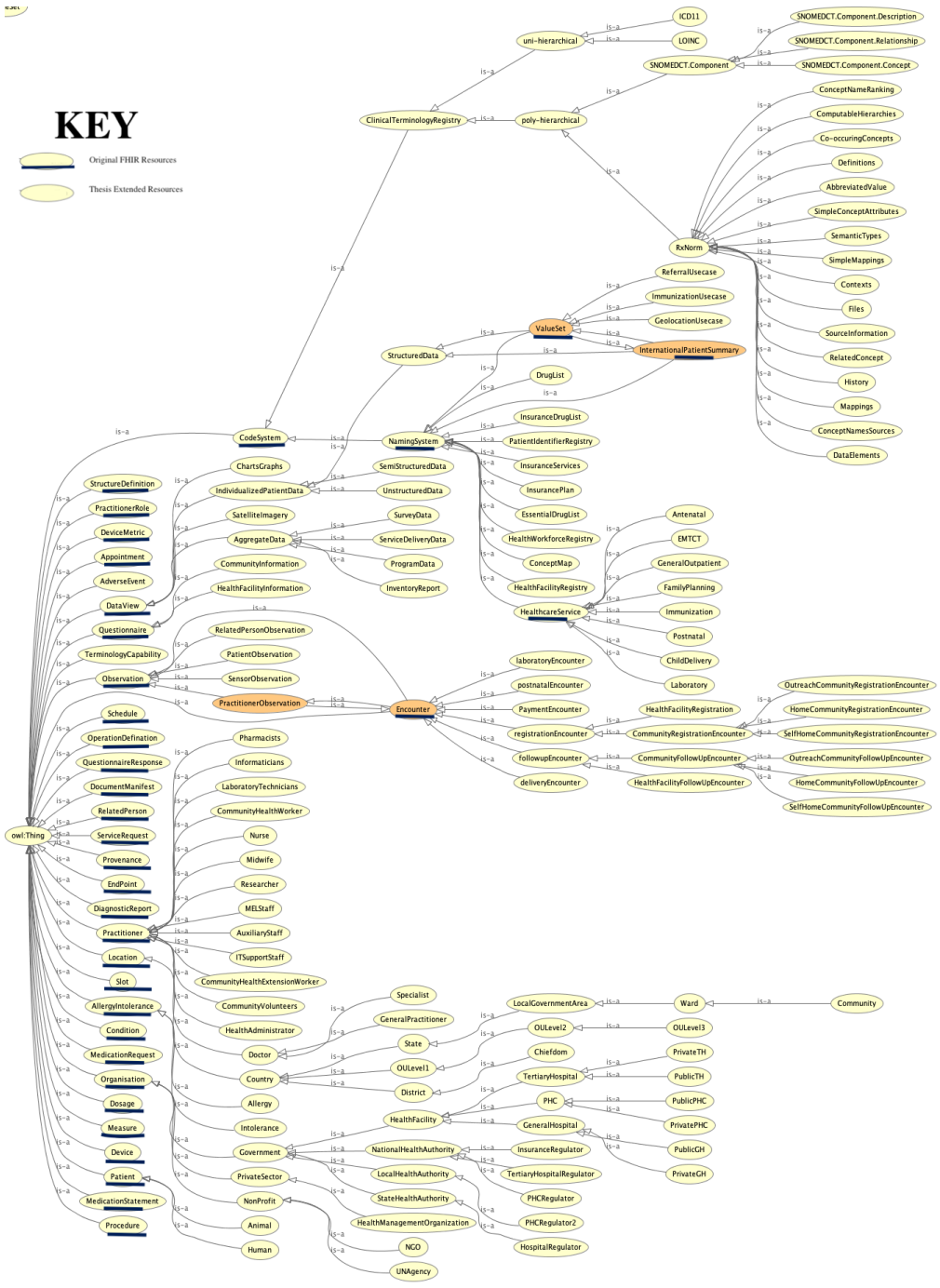


Figure A.3: RegistryChain detailed Ontology

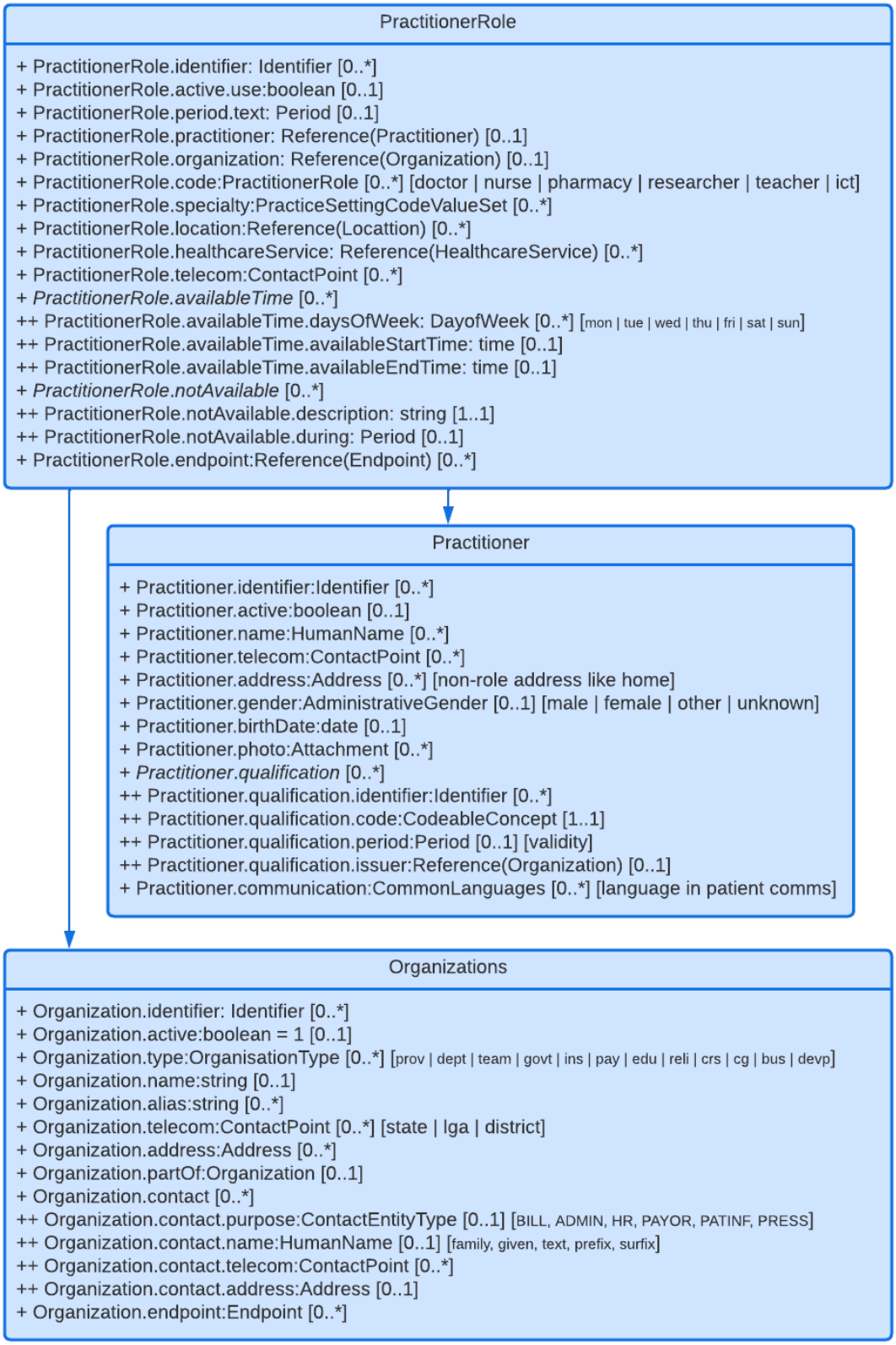


Figure A.4: The Practitioner, PractitionerRole, and organization FHIR resource data structure

Also, the Systematized Nomenclature of Medicines Clinical Terms (SNOMED CT) terminology attributes was also developed as a FHIR CodeableConcept model for validation. The SNOMED CT clinical concepts's class diagram is shown in Figure A.5

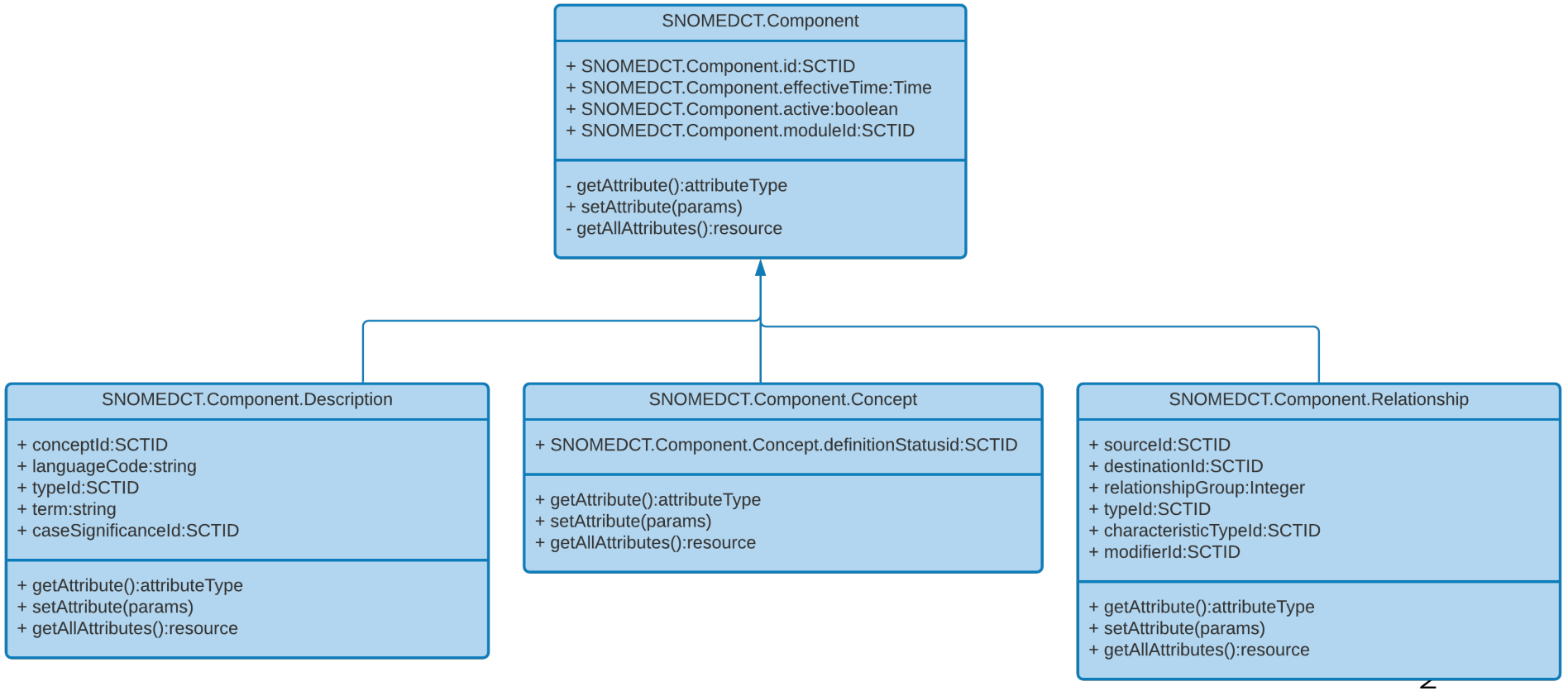


Figure A.5: SNOMED CT Class diagram

Similarly, health facility information was modeled using the organization FHIR resource in addition to Practitioner and PractitionerRole resources.

A.1.3 On-chain smart contract models

Smart contracts in fabric extend the $Contract$ class [253]. The $Contract$ class has several functions. All functions in a smart contract pass a transaction context object ctx (of $Context$ class) as first argument. The $ChaincodeStub$ class is implemented in smart contract as $stub$ object exposed by ctx. Amongst other operations, the $putState$ operation with arguments key and $value$ is used to add data to the blockchain world state. The Node.js (JavaScript) Software Development Kit (SDK) called *fabric-chaincode-node* was used to implement smart contract for the FHIR structured data and tokens. The *fabric-chaincode-api* and the *fabric-shim-api* node packages were used. Before a smart contract is used, it is packaged in a chaincode, the chaincode installed, approved, committed, and initialized as described later in next section.

A.2 Network Setup

As noted earlier, the experiment is implemented using the Hyperledger fabric blockchain. In order to set up the blockchain node for the reference implementation, the following steps were taken:

- Install pre-setup software and setup environment
- Defining configuration files
- Download Hyperledger Fabric source code from https://github.com/hyperledger/fabric
- extract the download to a convenient folder, change directory into the folder
- Download binary tools
- Generate cryptographic materials
- Generating Network materials
- Setup and start the Orderer
- Setup and start the Peer
- Create a channel and join the peer

A.2.1 Install pre-setup software and setup environment

The environment for Hyperledger fabric nodes (peers, orderers, and CAs) was setup using docker containers or hosted as processes. Both can be run on a virtual machine or host machine (which can be Ubuntu, Windows, or Mac). A single virtual machine or multiple virtual machines may be used for more component setup. In this case, multiple docker containers is used, one for each network component (peer, orderer, and CA). The prerequisites installed on our Ubuntu 20.0 machine include $build\text{-}essential, git, make, unzip$, g++, $libtool, jq, docker$, $docker\text{-}compose, go, npm, and node$. At runtime, the docker container for each network component will be initialized using the docker-compose files in the github repository.

The container for the peer will have access to the peer organization's msp cryptographic materials (certificates and keystore) folder in the physical or virtual machine. Similarly, there will be a different container for the orderer that will have access to the cryptographic materials in the orderer organization's msp folder (or tls folder if tls is enabled). If a process on the virtual machine is used (instead of a docker container), then the process will need access to the requisite cryptographic material.

A.2.2 Defining configuration files

The main configuration files are *crypto-config.yaml*, $configtx.yaml$, $orderer.yaml$, and $core.yaml$. The *crypto-config.yaml* is used to generate cryptographic identity materials (as needed), while the $configtx.yaml$ is used to generate the genesis block and the channel transaction file. Similarly, the $orderer.yaml$ configuration file is used to setup the orderer before launch. The $core.yaml$ is used for peer configuration. All properties within these files may be overwritten by environment variable like NAME_SECTION_SUBSECTION_PROPERTY e.g., CORE_PEER_ID. In this case, as indicated in the shell scripts in the source code, some of these attributes have been overwritten with the corresponding environment variables.

crypto-config.yaml The cryptographic materials are certificates and keystores for identification, encryption, decryption, and signing. This can alternatively be provided by a CA. The *crypto-config.yaml* file for generating the cryptographic material has two parts *OrdererOrgs* and *PeerOrgs*. The *OrdererOrgs* section has a list of orderers with main attributes as name and domain. The *PeerOrgs* section has list of Peer-organizations with name and

domain attributes amongst others. Under each Peer-organizations, a subsection *Template* has an attribute count that can be set to determine number of peers with template to create. Similarly, the *Users* subsection use the attribute count to set the number of users in addition to the default admin. The *PeerOrgs* section also have a CA subsection for implicit definition of organization CAs. Next, three crypto-config files for orgA, orgB, and orgC named *crypto-config-orgA.yaml*, *crypto-config-orgB.yaml*, and *crypto-config-orgC.yaml* used for the reference network were created.

configtx.yaml The $configtx.yaml$ file is used for network configuration, and it has six sections whose configuration affect the network structure and configuration. The sections are: $Organizations$, $Orderers$, $Applications$, $Channel$, $Capabilities$, and $Profile$. The $Organizations$ section lists all organizations with tags for easy reference in other parts of the $configtx.yaml$ file (eg. tag: &orgA and reference: - «: *orgA). Each organization tag will have attributes - Name, SkipAsForeign, ID, MSPDir, Policies, OrdererEndpoints, AnchorPeers. The capability defines the Hyperledger version for Channel, Orderer, and Application, all set to version two and above. The $Application$ section is tagged too, and contain Access Control Lists (ACLs) for different blockchain and chaincode functions. The $Application$ section also allow setting the list of participating organizations, in this case OrgA, OrgB, OrgC on the application side. The Policies subsection is where policies for LifeCycleEndorsement, Endorsement, Readers, Writers, and Admins are set. The $Orderer$ section has OrdererType which can be set to solo, kafaka, or etcdraft (code for raft). It also has a list of the three orderer addresses (orderer0.orgA.com, orderer0.orgB.com, orderer0.orgC.com), BatchTimeout, BatchSize, MaxChannels, EtcdRaft Consenters, and Options lists. The orderer section also has a list of ordering organizations (OrgA, OrgB, OrgC), Policies, and Capabilities. The $Channel$ section defines the Policies for Readers, Writers, and Admins: each having Type and Rule attribute. The $Profile$ section was used to define the Channel, Orderer, Application, and Consortium configuration to use for the Profile.

orderer.yaml The $orderer.yaml$ is used for documenting the configuration of the network orderer(s). In the reference network, there are three orderers orderer0.orgA.com, orderer0.orgB.com, orderer0.orgC.com, owned by each of the three organizations A, B, and C. The $orderer.yaml$ file has eight sections: General, FileLedger, Debug Configuration, Operations Configuration, Metrics Configuration, Admin Configuration, Channel participation API configuration, and Consensus configuration. The General section has the general properties for the Orderer. The FileLedger is the location of the file system where the ledger is located, and where the Orderer persists the blockchain ledger data. The properties of the ordererType etcdraft is included in the Consensus configuration section. The Operations endpoints exposed by Orderers are used for monitoring the Orderer and alerts

configuration. The Operations parameters are set in the Operations section. The Metrics configuration help set parameters of orderer emitted metrics that can be used by third party application for various other purposes.

The Cryptographic Service Provider (CSP) helps with all cryptographic functions - encryption, decryption, digest creation, hashing and more. The CSP can be implemented either software-based or hardware-based. The software CSP is platform specific libraries. The Hardware Security Module (HSM) are used for implementing CSP. The hardware CSP are more secure, and are based on the PKCS11 platform independent standard. The orderer can be configured to use either the software or hardware depending on the developer's preference. The parameters are set under the Blockchain Crypto Service Provider (BCSP) in the respective orderer.yaml file. The options under the $default$ attribute are SW or PKCS11 standing for software or hardware based CSP respectively. The SW subsection has to be set when SW is the option. The BCSP must be consistent across the network. In the reference network, the SW option was chosen.

The orderer will accept connections when listenAddress and port are set. Enabling tls on the orderer is important for client authentication. The orderer can be initialized with or without the system channel, initialized the test network with the system channel. In this case, the bootstrapMethod variable under General section in orderer.yaml file is set to $file$. The location of the $genesis.block$ file is then set using the variable BootstrapFile.

core.yaml The $core.yaml$ file is used to configure and manage the runtime behavior of the peer. The three $core.yaml$ files corresponding to the peer configurations one each for the three organizations were configured. $core.yaml$ file has several sections. The $peer$ section control the general behavior of the peer like network, MSP, storage path, private data behavior, authentication, and more. the $ledger$ section is used to set the type of the state database, which can be couchDB or levelDB. The $chaincode$ section control the runtime behavior of the chaincode. The $operations$ and $metrics$ sections are both used for managing the endpoint configuration for peer monitoring and alerting, through third party products. The vm section is used to configure the use of docker container in virtual machines.

A.2.3 Generate cryptographic materials

As described in Chapter four, the identity can be managed by an embedded Hyperledger fabric CA server. The server is initialized and launched respectively with the *fabric-ca-server init* and *fabric-ca-server start* commands. The full start command for example will

be *fabric-ca-server start -b admin:adminpw*. The server binary has several flags that can be used with commands for various purposed. For the purpose of testing this model, identities were generated locally using the *cryptogen* binary tool.

The *cryptogen* binary tool is CLI utility for generating crypto materials on test Hyperledger fabric network. The generated crypto materials are the identity certificates and the key stores. The developer (or admin) is expected to provide network participants (organizations, users, peers, and orderers) configuration information in *.yaml* file format (eg. crypto-config.yaml). The *cryptogen* tool generates the certificates and key stores in the local folder. Line 1 is first used to set the environment variable to the location for the *cryptogen* binary. A sample is then first generated from Line 2 and pipped to file *temp.yaml* and can be edited as needed. The file was edited to suite the configuration as in the git repository and use Line 3 to generate the cryptographic materials stored and grouped by organizations.

```
1    export PATH=$(pwd)/build/bin:$PATH
2    cryptogen showtemplate > temp.yaml
3    cryptogen generate
     --config=<<config file path>>
     --output=<<output folder>>
```

Folder structure named *Organizations*, which contains *OrdererOrganizations* and *PeerOrganizations* folders. The *OrdererOrganizations* sub-folder for instance will contain an *orderers* folder with a list of ordererFolders. In each orderer folder has an *msp* and a *tls* sub folders which contain folders *admincerts*, *cacerts*, *keystore*, *signcerts*, and *tlscerts*. These sub-folders each contain the Admin's public certificate, the Certificate of the CA provider, the orderer's private key, the orderer's public key, and the orderer's Transport Layer Security (TLS) certificate. As TLS is enabled, the *tls* folder contents was used instead. And access to the orderer *tls* sub folder is provided. This access is provided by setting the ORDERER_GENERAL_LOCALMSPDIR to the *tls* folder path.

```
4    export ORDERER_GENERAL_LOCALMSPDIR=
     $(pwd)/<<path to orderer msp folder>>
5    export ORDERER_GENERAL_LOCALMSPDIR=
     $(pwd)/<<path to orderer tls folder>>
```

The peer node follow a similar folder structure *peerOrganizations>OrgA>peers>peer0.orgA.com>msp/tls*.

```
6    export CORE_PEER_MSPCONFIGPATH=
     $(pwd)/<<path to peer msp folder>>
7    export CORE_PEER_MSPCONFIGPATH=
       $(pwd)/<<path to peer tls folder>>
```

The admin user created by default also has the *msp* and *tls* directories under the *users* subdirectory of say *peer0.orgA.com*. The administrator need to setup access to certificates to be able to configure changes like start or stop *orderers* and *peers*. The admin will also need to be able to install or start the *Chaincode* and make other network changes as necessary. Similarly, users will use the certificates under their respective folders under the *users* folder to Invoke and Query *Chaincodes* on *peer*.

In order to generate additional network participant(s)' cryptographic materials, the *extend* command is used instead of *generate* command. A separate *extension.yaml* file is created and used as *config* input to the *extend* command. This command was not used in this case. This command can be used to add additional *orderer*, *peers*, *organizations*, *users*, *admins*.

```
8    cryptogen extend --config=<<extension file path>>
     --input=<<existing_cryptogen_generated_folder>>
```

A.2.4 Generating Network material

The *configtxgen* tool is a command line utility for generating and managing network configuration artifacts. The artifacts are the GenesisBlock and the ChannelTransaction used on the Hyperledger fabric network. The developer (or admin) is expected to provide parameters and inputs (organizations, individual, peers, and orderers) configuration information in .yaml file format (using the *configtx.yaml* created in previous section). Generated files are in binary format and cannot be read in text editors, hence the *inspection* command is used to output the generated content in JSON format. The environment variable FABRIC_CFG_PATH is first set to the folder path containing the *configtx.yaml* file in Line 1. Line 2 of the command is used to generate the network **genesis block**. The generated file *genesis.block* is used for launching the *orderer*.

```
1    export FABRIC_CFG_PATH=
     $(pwd)/<<path_to_folder_with_configtx.yaml>>
2    configtxgen -outputBlock ./artifacts/genesis.block
      -profile RegistryChainProfile
      -channelId RegistryChainChannel
```

The Hyperledger fabric network does not have to be on for the commands to work. The content of *genesis.block* is inspected using line 3.

```
3    configtxgen -inspectBlock ./artifacts/genesis.block
      > ./artifacts/genesisJSON.json
```

Next, the **network configuration transaction** file is generated. Line 4 is used to generate the channel transaction file. The command $outputCreateChannelTx$ requires that $Application$ and $Consortium$ sections under profile used to be populated in the $configtx.yaml$. The consortium organization members' msp folders need to be accessible to the $configtxgen$ tool so that these members can assess the channels. The created $channel.tx$ is used by the $peer$ binary for the submission of transaction for the creation of application channel. The created $channel.tx$ binary file can be inspected in line 5 by piping to $channelJSON.json$ file.

```
1     4       configtxgen
2           -outputCreateChannelTx ./artifacts/channel.tx
3             -profile RegistryChainChannel
4             -channelId channelID
5     5       configtxgen
6           -inspectCreateChannelTx ./artifacts/channel.tx
7              > ./artifacts/channelJSON.json
```

Line 6 is used to print the information about any organization given the name and saved as json.

```
1     6       configtxgen -printOrg <<organization name>>
2             > ./artifacts/OrgName.json
```

A.2.5 Setup and start Orderer

The *Orderer binary* is used to order transactions in Hyperledger fabric and implemented as a single binary file. It takes the *genesis.block* for initialization. The runtime properties of the Order binary is managed in the $orderer.yaml$ configuration file. In this case, some of these properties were overwritten using the environment variable approach previously described. The messaging infrastructure for Hyperledger fabric can be implemented either as $solo$, $kafka$, or $raft$. The $raft$ ordering is recommended for Hyperledger fabric network, and the other two are being deprecated. As a result, the $raft$ ordering service was used for the test implementation. Raft is inbuilt into Hyperledger fabric, and supports clustering. See the component representation of the RegistryChain use case in Figure A.6. The Orderer binary needs access to the generated genesis block file (eg. genesis.block). At runtime, the orderer writes the details of the genesis block to the blockchain ledger file system. The location of the ledger is set in the $orderer.yaml$ configuration file. The orderer messaging service expose gRPC service to peers and clients. The gRPC use *Protocol Buffers* $protocobuf$. The security on orderer binary is facilitated by enabling TLS on the orderer binary. The orderer binary does not have an inbuilt cryptography manager,

it leverages on external crypto service provider for encryption, decryption, and message signing.

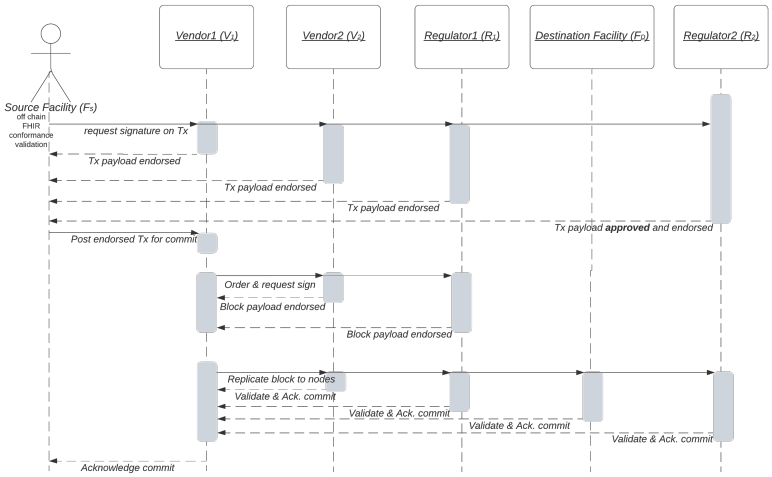


Figure A.6: Component diagram of Raft Ordering service

The first step after the download the fabric binaries is to execute Line 1 to build the $orderer$, $peer$, and $configtxgen$ binaries in a new folder *build/bin* subdirectory. Line 2 of the code snippet is then used to place the subdirectory in the environment path. Line 3 is used to set the FABRIC_CFG_PATH environment variable to the path of the directory that contain sample configurations included in the downloaded Hyperledger fabric source files. The orderer binary is then given access to the orderer msp folder.

```
1    make orderer peer configtxgen
2    export PATH=$(pwd)/build/bin:$PATH
3    export FABRIC_CFG_PATH=
       $(pwd)/<<path_to_folder_with_orderer.yaml>>
```

After the initialization of the orderer using the genesis bock file, it can then be launched. Next, steps before starting the orderer, the *hyperledger* subdirectory in the *var* subdirectory was created, and then assigned ownership to the device *user_name*. The $/var/hyperledger$ is the default directory for storing Hyperledger fabric ledger data by levelDB from docker containers. This may be changed as desirable. This change was not effected in this case.

```
4    sudo mkdir /var/hyperledger
5     sudo chown user_name /var/hyperledger
```

The orderer is then started with the code snippet in line 6. After the command is fully, executed, the orderer will be up and running with the Command Line Interface (CLI) interface showing the output *[orderer.common.server] Main -> Beginning to serve requests*. This

will lead to the orderer successfully now managing the blockchain ledger data in the file system. The orderer will need four things already available, i) *access to generated crypto material*, ii) *genesis block file*, iii) *channel transaction file*, and iv) *access to ledger folder*. The orderer will write the ledger data to the provided ledger folder. This location is set in the $Location$ element under the $FileLedger$ section of the $orderer.yaml$ file.

```
1    6    orderer
```

A.2.6 Setup and start Peer

The peer binary is used to launch as a process to become a Hyperledger node on a blockchain network. A peer binary can also serve as utility for managing network configuration or the network peers. The peers are configured using the $core.yaml$ file. The peers were configured by navigating to the directory with downloaded, and extracted folder and add PATH and FABRIC_CFG_PATH to the environment path with Lines 1 and 2. Line 2 is used to set the FABRIC_CFG_PATH to the folder containing the $core.yaml$ file. The peer binary also have an embedded gRPC server that exposes a gRPC service to connect with clients, applications, other peers, and orderers to communicate with the peer. Peer leverages the CSP. As tls is enabled, then the host name is set with peer name by editing the $/etc/hosts$ file.

```
1    1     export PATH=$(pwd)/build/bin:$PATH
2    2     export FABRIC_CFG_PATH=$(pwd)
3          /<<location_of_core.yaml>>
```

The $peer$ command of the peer node is used to manage the node, channel, or chaincode in the form:

```
1    3     peer <<command>>
```

The «commands» are $node$, $channel$, and $lifecycle$ for managing them respectively. the $lifecycle$ command has the «subcommand» $chaincode$. The orderer must be up for the peer to function. The $peer\ node\ start$ command starts the peer.
Next, the peer is started by executing the command in line 4 (if error is thrown about port being in use, change the port in *core.yaml->Operations->listenAddress* to another port.

```
1    4     peer node start
```

A.2.7 Create a channel and join peers

The peer command can be used with *channel* command to *create*, *join*, *list*, *fetch*, or *update* channels. Similarly, both can be use with *signconfigtx* to sign channel transaction. See how the transaction file $tx_file.tx$ is signed in Line 3.

```
1    export PATH=$(pwd)/build/bin:$PATH
2    export FABRIC_CFG_PATH=$(pwd)
     /<<path to createChannelTx_file>>
3    peer channel signconfigtx -f <<tx_file.tx>>
```

The peer uses the information in the *core.yaml* file to create the channel by initiating a gRPC communication with the orderer to create the channel. If the create invocation is successful, a file with name *channelID.block*, which is the genesis block is written to the peer file system. Other peers joining the network must have access to the genesis block. The actual channel creation is executed in Line 6.

```
4    export PATH=$(pwd)/build/bin:$PATH
5    export FABRIC_CFG_PATH=$(pwd)
     /<<location_of_core.yaml>>
6    peer channel create -o ordererUrl
       -c channelID -f <<signed_channel_tx_file>>
```

Other peers will receive channel blocks through a *fetch* operation. The join operation is executed using the admin identity of joining organizations. The successful join will result in initialization of the ledger and state database of the channel on the joining peer. The number 0 is set for genesis block, and the command *config* is used to retrieve the latest configuration block for the network. The commands are executed on the orderer and has to be set as per of the command.

```
7    export PATH=$(pwd)/build/bin:$PATH
8    export FABRIC_CFG_PATH=$(pwd)
     /<<location_of_core.yaml>>
9    peer channel fetch 0 -o ordererUrl
       -c channelID <<outputFileName>>
10   peer channel join -o ordererUrl
       -b <<genesis_block_file>>
```

A.2.8 Signatures for network artifacts

The administrator of one organization will fetch the network config file from the orderer using the command in Line 13.

```
11      export PATH=$(pwd)/build/bin:$PATH
12      export FABRIC_CFG_PATH=$(pwd)
        /<<location_of_core.yaml>>
13      peer channel fetch config -o ordererUrl
        -c channelID <<outputFileName>>
```

The admin will then use the *configxlator* binary to make changes to the network transaction file, and then the required number of admins will sign using the *peer channel signconfigtx*. No new file is generated when the file is signed (add admin certificate, and sign file with admin private key). The size changes at the end of the command.

```
14      peer channel signconfigtx -f <<tx_file>>
```

One of the admins will then submit the transaction to the network using *peer channel update* command.

```
15    peer channel update
        -f <<signed tx_file>>
        -c channelID
        -o ordererUrl
```

A.2.9 Chaincode operations

Packaging chaincode The *peer* Command Line Interface (CLI) command is used to create a packaged chaincode file *RegistryChain.tar.gz* using Line 1.

```
1  peer lifecycle chaincode package RegistryChain.tar.gz
 --path ../path/to/SmartContractFile
 --lang node
 --label RegistryChain_v1
```

Installing the Chaincode Packaged smart contracts can be installed by each of the collaborating organizations. The packaged chaincode is installed on organization's endorsing peers (each organization's only peer). The command is used for the installation step.

```
1 peer lifecycle chaincode install RegistryChain.tar.gz
```

Installed chaincode can be verified by executing the command to get the details of the installed chaincode.

```
2  peer lifecycle chaincode queryinstalled
```

The output will be the list of *package-ids, eg. xx335de39xx* for installed chaincode packages in the environment.

Package approval Depending on the endorsement policy, a set number of organizations will have to approve the installed chaincode package before it is accessible on the

network channel. The *package-id* is used to uniquely identify installed packages.

```
  peer lifecycle chaincode approveformyorg -o localhost:7050
    --ordererTLSHostnameOverride orderer0.orgA.com
    --tls true
    --cafile /.../tlsca.orgA.com-cert.pem
    --channelID RegistryChainChannel --name registrychain
    --version 1 --init-required
    --package-id registrychainchannel_v1:xx335de39xx
    --sequence 1
```

Commit chaincode definition The *checkcommitreadiness* command can be used to check if the chaincode definition can be committed. The number of organizations is determined by chaincode endorsement policy. The life-cycle endorsement policy can be either ANY, $MAJORITY$, or ALL. The $MAJORITY$ *Endorsement* is the default in Hyperledger Fabric. In this case, the $MAJORITY$ endorsement is set. The command will return a json of organizations' approvals.

```
peer lifecycle chaincode checkcommitreadiness
    --channelID RegistryChainChannel --name registrychain
    --version 1 --init-required
    --sequence 1
    --output json
```

The chaincode definition can be committed after obtaining necessary approvals using the command.

```
peer lifecycle chaincode commit
    -o localhost:7050
    --ordererTLSHostnameOverride orderer0.orgA.com
    --tls true
    --cafile /.../orderers/tlscacerts/tlsca.orgA.com-cert.pem
    --channelID RegistryChainChannel
    --name registrychain
    --peerAddresses localhost:7051
    --tlsRootCertFiles /.../peer0.orgA.com/tls/ca.crt
    --version 1
    --sequence 1
    --init-required
```

The next step is to query and check that the chaincode is committed using the command *querycommitted* chaincode command. The output will tell its readiness for operations.

```
peer lifecycle chaincode querycommitted
    --channelID RegistryChainChannel --name registrychain
```

Initialize smart contract The smart contract is initialized using the *invoke* chaincode command. Executing the command will output the status response of 200 if successful.

```
peer lifecycle chaincode invoke
    -o localhost:7050
    --ordererTLSHostnameOverride orderer0.orgA.com
    --tls true
    --cafile /.../orderers/tlscacerts/tlsca.orgA.com-cert.pem
    --channelID registryChainChannel
    --name registrychain
    --peerAddresses localhost:7051
    --tlsRootCertFiles /.../peer0.orgA.com/tls/ca.crt
    --peerAddresses localhost:9051
    --tlsRootCertFiles /.../peer0.orgB.com/tls/ca.crt
    --isInit
    -c '{"function":"initLedger", "Args":[]}'
    --init-required
```

The smart contract is now ready to serve clients. A query can be executed as

```
peer chaincode query
    --channelID RegistryChainChannel
    --name registrychain
    -c '{"Args":["queryAllOrganizations"]}'
```

A.2.10 Security and Privacy

The Security and privacy of data in the world state and on the ledger can be managed using one or more combination of: $1)$ Client side data encryption, $2)$ On chain Private Data Collection (PDC), $3)$ Attribute based access control and $4)$ Channel segmentation with organization MSP managed Identities.

Similarly, the security of the peer is enhanced by setting the CORE_PEER_ADDRESS property in the *core.yaml* configuration file for peer configuration to internal organization address. This will ensure that CLI access to the peer is only from the internal network.

Clients can use any encryption algorithm of choice alongside other measures to secure data stored on the blockchain both in transit and at rest. Similarly, Private Data Collection (PDC), a partition on the shared ledger dedicated to an organization for information isolation even on same channel. Private data of an organization can be shared, it can be set to expire after a certain number of blocks or on demand. Attribute based access control gives more granularity for authentication and authorization to implementers. The

access control depends on the client and application identity. The identities of network participants belong to organizations who then belong to channels on the blockchain infrastructure.

A.2.11 Maintenance

The main maintenance tasks are organization addition, node addition, and user addition. All these can be conducted using fabric CA or cryptogen binary.

A.2.12 Peer events

Peers emit events on receiving blocks from orderer. Each other peer that receive the blocks through gossip protocol also emit events. Clients then subscribe to these events. Events subscription is local to each channel (only member organizations in a channel can subscribe). Events can be filtered or un-filtered. In filtered events, only transaction summary, status data, and event name are included. In the un-filtered mode, the event payload including all transaction information is included in the event.

A.2.13 Distributed Application

The thesis implemented the application in Node.js (JavaScript) and plan to make this available as open source global good when mature as a node SDK.