Blockchain Adoption in Healthcare : Toward a Patient Centric Ecosystem

Rita Azzi

INSTITUT POLYTECHNIQUE DE PARIS

NNT : 2023IPPAT053

Thèse de doctorat

Blockchain Adoption in Healthcare: Toward a Patient Centric Ecosystem

Thèse de doctorat de l'Institut Polytechnique de Paris
préparée à Télécom Paris en cotutelle avec l'Université Saint-Joseph de Beyrouth

École doctorale n°626 École doctorale de l'Institut Polytechnique de Paris (ED IP Paris)
Spécialité de doctorat : Informatique

Thèse présentée et soutenue à Palaiseau, le 22/12/2023, par

RITA AZZI

Composition du Jury :

Dominique Gaiti Professeure, Université de Technologie de Troyes, France	Présidente/Examinatrice
Pascal Lorenz Professeur, Université de Haute-Alsace, France	Rapporteur
Bijan Jabbari Professeur, George Mason University, Washington, United States	Rapporteur
Farid Naït-Abdesselam Professeur, Université Paris Cité, France	Examinateur
Rida Khatoun Professeur, Télécom Paris, France	Examinateur
Ahmed Serhrouchni Professeur émérite, Télécom Paris, France	Directeur de thèse
Rima Kilany Chamoun Professeure, Université Saint-Joseph de Beyrouth, Liban	Co-directrice de thèse
Maria Sokhn Professeure, Haute École Spécialisée de Suisse Occidentale, Suisse	Co-encadrante/ Invitée

On dit qu'avant d'entrer dans la mer,
une rivière tremble de peur.
Elle regarde en arrière le chemin
qu'elle a parcouru, depuis les sommets,
les montagnes, la longue route sinueuse
qui traverse des forêts et des villages,
et voit devant elle un océan si vaste
qu'y pénétrer ne parait rien d'autre
que devoir disparaître à jamais.
Mais il n'y a pas d'autre moyen.
La rivière ne peut pas revenir en arrière.
Personne ne peut revenir en arrière.
Revenir en arrière est impossible dans l'existence.
La rivière a besoin de prendre le risque
et d'entrer dans l'océan.
Ce n'est qu'en entrant dans l'océan
que la peur disparaîtra,
parce que c'est alors seulement
que la rivière saura qu'il ne s'agit pas
de disparaître dans l'océan,
mais de devenir océan.

La peur - Gibran Khalil Gibran -

Acknowledgements

James Norbury, once wrote “Which is more important? [. . .] The journey or the destination? [It’s] The Company”. I am profoundly grateful for the exceptional support and companionship I received throughout my PhD journey. I will forever be indebted to those who supported me every step of the way.

First, I would like to thank God for giving me the strength to persevere in my PhD journey. He was my rock throughout all the challenging moments of completing this thesis.

I would like to express my gratitude to my Ph.D. advisors, Professor Ahmed Serhrouchni and Professor Rima Kilany Chamoun, for their intangible support and invaluable knowledge shared with me. To Professor Ahmed Serhrouchni, I am thankful for your dedication to fostering my research skills and consistently challenging me to think critically. You not only shared your knowledge but also your life philosophy, shifting our discussion from academic topics to everyday life with a blend of seriousness and humor. To Professor Rima Kilany Chamoun, I am thankful for your unwavering presence, insightful feedback, and guidance throughout my journey. You encouraged me and gave me incredible opportunities to engage in student projects. These experiences broadened my horizons and challenged me to grow and develop, striving for excellence. This endeavor would not have been possible without the generous support of Professor Maria Sokhn, who has influenced and inspired me. Her precious advice and observations motivated me to progress as a researcher.

I would like to thank the National Council for Scientific Research of Lebanon (CNRS-L) for awarding me a doctoral fellowship and Saint-Joseph University for granting me a tuition waiver.

Besides, I would like to thank all the jury members: Prof. Dominique Gaiti, Prof. Pascal Lorenz, Prof. Bijan Jabbari, Prof. Rida Khatoun and Prof. Farid Naït-Abdesselam for their interest in my work and their insightful comments and suggestions.

I extend my gratitude to my friends who have become like a second family to me, including those who supported me before embarking on my PhD journey and those I've met along the way.

Last but not least, my deepest appreciation goes to my father Elie, my mother Marie-Thérèse and to my sister Cynthia. Without their support and their prayers, I would not have had the strength to complete this work. Their faith and trust in me kept me motivated throughout the entire journey.

Abstract

The healthcare sector evolves constantly, driven by technological advancement and innovative solutions. From remote patient monitoring to the Internet of Things (IoT), Artificial Intelligence (AI), personalized medicine, mobile health, and electronic records systems, technology has improved patient outcomes and enhanced care delivery. These technologies have shifted the healthcare ecosystem to be more patient-centered, focusing on meeting the patient's needs rather than the needs of the individual organizations within it. However, this transformative shift experienced by the healthcare industry is associated with multiple challenges due to the inherent complexity and fragmentation of the healthcare ecosystem. This thesis addresses three healthcare ecosystem challenges that significantly impact patients. The first challenge addressed is the problem of counterfeit or falsified drugs that represent a threat to public health, resulting from the vulnerabilities in the pharmaceutical supply chain, notably centralized data management and the lack of transparency. The second challenge addressed is the problem of healthcare data fragmentation that thwarts care coordination and impacts clinical efficiency. This problem results from the dynamic and complex patients' journey in the healthcare system, shaped by their unique health needs and preferences. Patient data are scattered across multiple healthcare organizations within centralized databases and are ruled by policies that hinder data sharing and patients' empowerment over their data. The third challenge addressed is the confidentiality and privacy of healthcare data that, if compromised, shatter the trust relationship between patients and healthcare stakeholders. This challenge results from the healthcare organizations' poor data governance that increases the risk of data breaches and unauthorized access to patient information.

The blockchain has emerged as a promising solution to address these critical challenges. It was introduced into the healthcare ecosystem with the promise of enforcing transparency, authentication,

security, and trustworthiness. Through comprehensive analysis and case studies, this thesis assesses the opportunities and addresses the challenges of adopting the blockchain in the healthcare industry. We start with a thorough review of the state of the art covering the blockchain's role in improving supply chain management and enhancing the healthcare delivery chain. Second, we combine theoretical and real-world application studies to develop a guideline that outlines the requirements for building a blockchain-based supply chain. Third, we propose a patient-centric framework that combines blockchain technology with Semantic technologies to help patients manage their health data. Our fourth contribution presents a novel approach to data governance by developing a blockchain-based framework that improves data security and empowers patients to participate actively in their healthcare decisions. In this final contribution, we widen the scope of the proposed framework to include a roadmap for its adoption across diverse domains (banking, education, transportation and logistics, etc.).

Contents

List of Tables

List of Figures

List of Acronyms

ADE	Adverse Drug Event
AI	Artificial Intelligence
API	Application Programming Interface
AVISPA	Automated Validation of Internet Security Protocols and Application
BFT	Byzantine Fault Tolerance
BLE	Bluetooth Low Energy
CA	Certificate Authority
CDS	Clinical Decision Support
CDSS	Clinical Decision Support System
CFT	Crash Fault Tolerance
CID	Content Identifier
CL-AtSe	Constraint-Logic-based Attack Searcher
DAG	Directed Acyclic Graphs
DHT	Distributed Hash Table
EHR	Electronic Health Record
EMR	Electronic Medical Record
EOA	Externally Owned Account
ePHI	Electronic Protected Health Information
EPR	Electronic Patient Record
ERP	Enterprise Resource Planning
epSOS	Smart Open Services for European Patient
FDA	Food and Drug Administration
FMA	Foundation Model of Anatomy
GDBMS	Graph Database Management System
GDP	Good Distribution Practice
GDPR	General Data Protection Regulation
GPRS	General Packet Radio Service
GPS	Global Positioning System

HIPAA Health Insurance Portability and Accountability Act
HLPSL High-Level Protocol Specification Language
HTTP HyperText Transfer Protocol
ICD International Classification of Diseases
IPFS InterPlanetary File System
IoMT Internet of Medical Things
IoT Internet of Things
JSON JavaScript Object Notation
LOINC Logical Observation Identifiers Names and Codes
MAC Media Access Control
MEC Multi-access Edge Computing
MSP Membership Service Provider
NFC Near-Field Communication
NIST National Institute of Standards and Technology
OFMC On-the-Fly Model Checker
OWL Web Ontology Language
PBFT Practical Byzantine Fault Tolerance
PDC Private Data Collection
PHD Personal Health Data
PHI Protected Health Information
PHR Personal Health Record
PoS Proof of Stake
PoW Proof of Work
QR code Quick-Response Code
RDF Resource Data Framework
RDFS Resource Description Framework Schema
RFID Radio Frequency Identification
SATMC SAT-based Model-Checker
SMD Semantic Medical Devices
SNOMED Systemized Nomenclature of Medicine
SSN Semantic Sensor Network
SW Semantic Web
SWRL Semantic Web Rule Language
SWT Semantic Web Technologies
TA4SP Tree Automata-based Protocol Analyzer
3G Third Generation
UMLS Unified Medical Language System
URL Uniform Resource Locator
URI Uniform Resource Identifier
WHO World Health Organization

Résumé Détaillé de la Thèse

Introduction

Tout comme une chaîne d'approvisionnement traditionnelle, l'écosystème de soins de santé est un système interconnecté composé d'organisations, d'individus, de processus, de ressources et de technologies impliqués dans la production et la fourniture de biens et services aux clients. Il est composé essentiellement des [1]:

- Fournisseurs de soins de santé : ils comprennent les professionnels de santé (tels que médecins, infirmières, etc.) et les établissements de santé autorisés à fournir des diagnostics ou des traitements médicaux (tels que les hôpitaux, les cliniques, les laboratoires médicaux, etc.).
- Industriels de santé : ils regroupent les fabricants de médicaments et de dispositif médicaux ainsi que les fournisseurs de logiciels de soins de santé.
- Financeurs de soins : ils comprennent principalement le gouvernement et/ou les compagnies d'assurance qui sont tous deux impliqués dans le financement des soins de santé.
- Laboratoires de recherche et de développement : ils comprennent des institutions publiques et privées qui participent au développement de nouveaux traitements médicaux et de technologies innovantes.

Ainsi, l'écosystème de soins de santé est un environnement dynamique, enchevêtré d'opportunités et de défis [2]. Il évolue constamment, sous l'influence des avancées technologiques qui l'ont orienté vers des approches centrées sur le patient, priorisant leur besoin qu'à ceux des différentes organisations qui le

composent. De même, le volume important de données collectées offre aux diverses parties prenantes de l'écosystème de santé de nouvelles perspectives puisque leur exploitation est indispensable aux innovations médicales. Toutefois, cette transition transformative que connait l'écosystème de soins de santé est associé à de multiples défis. Le patient est au sein d'un réseau complexe où chaque interaction entre les différents acteurs est prône à des erreurs. Les données peuvent être collectées, enregistrées ou partagées de manière inexacte ou incomplète. Les défaillances techniques qui peuvent perturber les services de soins et impacter la fourniture des médicaments ou des équipements médicaux. On cite notamment les pannes de courant, le dysfonctionnent d'un composant informatique, le plantage inattendu d'un logiciel. De même, les failles de sécurité informatique qui peuvent nuire considérablement à la relation de confiance entre les différents acteurs du système de santé. Il est important de noter que le secteur de la santé est également soumis à un ensemble complexe de normes et de règlementations qui peuvent entraver l'adoption rapide de solutions innovantes.

Puisque l'écosystème de soins de santé est un réseau de chaînes interconnectées, il est impératif d'acquérir une compréhension approfondie de ses dynamiques et de ses interactions afin de pouvoir relever les divers défis auxquelles il est confronté. Il est ainsi essentiel de dissocier la chaîne d'approvisionnement des soins de santé de la chaîne de prestation de soins pour une meilleur évaluation des défis potentiels, en vue de les résoudre. La chaîne d'approvisionnement des soins de santé englobe le mouvement des produits de santé depuis leur fabrication jusqu'à leur distribution aux différents acteurs de la chaîne. Elle est constituée de deux flux principaux : le flux physique et le flux d'information [2], [3].Une bonne gestion de la chaîne est essentielle pour assurer que les bons produits soient disponibles au bon moment et au bon endroit pour garantir la continuité des soins de santé. Cela implique la gestion des activités nécessaire pour garantir la qualité et la sécurité des produits tout au long de la chaîne ainsi que la gestion des données relatives à ces activités [2], [3]. La chaîne de prestation de soins englobe le mouvement des patients au sein du système de santé depuis leurs recours aux soins jusqu'à leurs rétablissements complet. Elle est constituée de deux flux principaux : le flux de patients et le flux d'informations [1], [3].Une bonne gestion de la chaîne est essentielle pour améliorer la qualité du parcours de soins du patient. Cela implique la gestion des activités, allant de la consultation médicale au diagnostic, traitement, suivi médicale, etc. Il s'agit également de gérer les données relatives à ces activités, notamment les données de santé [1], [3]. De plus, chacune de ces deux chaînes est caractérisée par un troisième flux : le flux financier. Ce flux fait référence aux opérations financières ayant lieu entre les parties prenantes impliquées dans ces chaînes. Il

inclut notamment, l'achat de fournitures médicales et le financement des activités de santé [2], [3].

Par conséquent, une quantité massive de données est générée quotidiennement dans l'écosystème de soins de santé, à la fois structurées et non structurées. Une bonne gestion de données est nécessaire pour maintenir la confiance entre les différents acteurs au niveau des chaînes d'approvisionnement et de prestation de soins de santé. Cet objectif est atteint grâce à une politique axée sur la transparence de la chaîne, assurant la disponibilité d'informations pertinentes au bon moment afin d'identifier rapidement les problèmes et aider à la prise de décision adéquate. Une bonne gestion de données nécessite alors, une compréhension approfondie des risques associés aux données de santé. Parmi les divers facteurs qui compromettent la sécurité des données et la vie privée des patients, on cite [1], [4]–[6]:

- La gestion centralisée des données qui constitue un point de défaillance unique qui compromet le contrôle et la souveraineté des données;
- Les menaces internes;
- Une mauvaise sécurité des données.

Cette thèse présente une problématique à trois niveaux. Chaque niveau traite un défi de l'écosystème de soins de santé ayant une répercussion sur la santé des patients ou portant atteinte à leurs vies privées. Le premier défi abordé est celui des médicaments contrefaits ou falsifiés qui représentent une menace pour la santé publique. Le deuxième défi concerne la fragmentation des données de santé qui entrave la coordination des soins et nuit à l'efficacité clinique. Le troisième défi s'attaque à la confidentialité des données relatives aux patients, ce qui implique aussi le respect du droit à la protection de la vie privée des patients. Une compromission de données ébranle la relation de confiance entre les patients et les acteurs du secteur de santé.

La blockchain apparait comme une technologie prometteuse, capable de relever ces différents défis [7]. Introduite dans l'écosystème de santé, la blockchain a le potentiel de renforcer la transparence, l'authentification, la sécurité et la fiabilité [8]. Néanmoins, cette technologie s'accompagne également de son lot de défis. Cette thèse évalue les risques et opportunités liés à l'adoption de la blockchain dans l'écosystème de soins de santé.

Pour répondre aux trois défis énoncés, nous proposons trois contributions. La première est une étude approfondie sur le rôle de la blockchain à améliorer la gestion de la chaîne d'approvisionnement. Pour compléter cette approche théorique, nous intégrons des applications concrètes du monde réel afin d'élaborer les exigences nécessaires à établir une chaine d'approvisionnement basée sur la blockchain. Cette contribution est résumée dans la section S1. Notre deuxième contribution, présente une approche axée sur le patient, où nous combinons la technologie blockchain et les technologies du Web sémantique pour aider les patients à gérer leurs données de santé. Cette contribution est présentée dans la section S2. Notre troisième contribution s'inscrit dans le cadre de la gouvernance des données. Nous développons un framework basé sur la blockchain pour améliorer la sécurité des données et qui par la suite pourra être adopter dans divers domaines. Ces travaux sont résumés dans la section S3.

S1- Défis de la Chaîne d'Approvisionnement des Soins de Santé : Parcours des Médicaments

La chaîne d'approvisionnement des soins de santé est un système complexe qui englobe tous les processus impliqués dans la conception, la fabrication et la distribution des produits de santé du fournisseur aux clients [9]. Elle est conçue pour maintenir la qualité des produits pendant toute la durée de leur transport à travers la mise en place de processus et de contrôles rigoureux pour garantir leurs conformités aux normes internationales. Certaines réglementations sont établies pour protéger les droits du consommateur. Parmi les huit droits fondamentaux du consommateur reconnus par les Nations Unies, on cite le droit à la sécurité, le droit à l'information, le droit de choisir et le droit à la réparation [10]. Aux États-Unis, la Food and Drug Administration (FDA) défend les droits des consommateurs en promouvant et en protégeant la santé publique, par le biais du contrôle et de la surveillance de la sécurité des produits alimentaires et produits de santé [11]. Plusieurs rappels de produits et alertes de sécurités ont été reportés par la FDA, montrant ainsi les nombreuses violations de la chaîne d'approvisionnement [12]. Ces incidents remettent en question la fiabilité de la chaîne d'approvisionnement et l'exactitude des données relatives aux produits. Les principaux objectifs de la chaîne d'approvisionnement ne sont pas pleinement atteints [13]. Notamment l'optimisation des coûts opérationnels, l'assurance de la qualité du produit conformément aux normes établies, la réduction des risques, la rapidité, la fiabilité, la durabilité et la flexibilité [13]. Non seulement

la santé du consommateur est touchée, mais les entreprises subissent aussi un préjudice moral qui nuit à leur réputation aux côtés des pertes financières. Dans le cas d'une chaîne d'approvisionnement pharmaceutique, la contrefaçon des médicaments et la gestion de la chaîne logistique du froid constituent l'un des principaux défis auxquels elle est confrontée. Ainsi, le principal risque de la chaîne d'approvisionnement se situe au niveau du parcours des médicaments. Ces risques résultent des vulnérabilités de la chaîne d'approvisionnement pharmaceutique, telles que la gestion centralisée des données et le manque de transparence de la chaîne, ce qui rend le système de traçabilité vulnérable. Le système est exposé alors à la corruption [14]. Le consommateur n'est pas en mesure de vérifier l'intégrité et l'authenticité des produits acquis. Nous avons besoin d'acquérir davantage de connaissances sur le produit, son origine, processus de fabrication et conditions de transport.

Un bon système de traçabilité vise à minimiser la production et la distribution de produits toxiques, contrefaits ou de qualité sous standard en améliorant l'étiquetage et le système de suivi. Les principaux composants d'un système de suivi sont les tags, traceurs et capteurs. Un tag est une étiquette placée sur le dessus d'un produit qui permet de l'identifier et de fournir des informations utiles relative au produit. Les tags les plus répandus sont : les codes QR (QR code), les codes-barres et les radio-étiquettes (RFID tags). Un traceur ou un marqueur est une substance ajoutée à un produit ou faisant partie de ces caractéristiques naturelles, utilisée pour fournir des informations sur le parcours et les processus que le produit a subi. Il est utilisé pour garantir que le produit n'a pas été altéré, certifiant ainsi son authenticité. Le capteur quant à lui est un dispositif qui permet de détecter les changements environnementaux tels que la lumière, la température, la pression et le mouvement. Les évènements détectés sont ensuite transmis à un dispositif électronique pour le traitement. Cependant, ces dispositifs de suivi sont parfois compromis et sujets au clonage. Par exemple, la technologie RFID a été intégrée dans la chaîne d'approvisionnement pour lutter contre les contrefaçons. Les acteurs de la chaîne d'approvisionnement peuvent suivre le parcours des produits à partir des informations transmises par les lecteurs RFID installés sur chaque produit. Toutefois, cette approche est vulnérable aux clonages [15]. Des produits falsifiés associés à des étiquettes clonées induisent le client en erreur et porte atteinte à sa santé [15]. Par conséquent, il est nécessaire de sécuriser les étiquettes et les dispositifs de suivi en garantissant leurs authenticités et unicités.

Outre les défis liés au système de suivi, nous devons s'attaquer au système de gestion centralisé de la chaîne d'approvisionnent, notamment le progiciel de gestion intégré (ERP) qui est exposé aux risques de cybersécurité [16]. Ces risques sont principalement liés à la base de données centralisées du ERP et aux modalités d'authentification peu robustes. Par conséquent, il est nécessaire de sécuriser les données col-

lectées et partagées avec tous les membres de la chaine d'approvisionnement et d'assurer leur disponibilité à la bonne personne et au bon moment.
La blockchain a été introduite dans la chaîne d'approvisionnement pour la rendre plus transparente, plus authentique et plus fiable [17]. La blockchain fournit un enregistrement immuable de toutes les transactions sur le réseau et réduit la nécessité des intermédiaires. Tous les détails relatifs aux produits ainsi que les informations relatives à leurs expéditions sont collectés par le biais de différentes technologies et validés avant d'être enregistrer de façon permanente sur la blockchain. L'objectif de ce travail est d'améliorer la traçabilité et l'authenticité des produits (notamment les médicaments) en intégrant la blockchain dans la chaîne d'approvisionnement pharmaceutique. Pour atteindre cet objectif, nous avons abordé trois questions de recherche :

- Quels sont les avantages de l'introduction de la blockchain dans la chaîne d'approvisionnement ?
- Est-ce que les informations partagées dans le système de traçabilité de la chaîne d'approvisionnent sont fiables ?
- Quels sont les défis à relever lors de l'intégration de la blockchain au sein de la chaîne d'approvisionnement ?

Afin de répondre à ces questions de recherche, nous avons adopté la stratégie des études de cas. Nous avons combiné des études théoriques et des études de cas pratiques pour élaborer notre théorie sur les exigences nécessaires pour améliorer la transparence de la chaîne d'approvisionnement [18]. Plusieurs startups ont déjà identifié la blockchain comme un nouveau paradigme visant à améliorer la visibilité et la gestion de la chaîne d'approvisionnement. Nous avons sélectionné Ambrosus [19] et Modum [20], deux startups suisses, comme cas concrets à étudier, car nous avons pu obtenir suffisamment de détails techniques sur leurs systèmes. Modum se spécialise dans la chaîne d'approvisionnement pharmaceutique et travaille à assurer la livraison des médicaments conformément aux exigences des bonnes pratiques de distribution (GDP). Ambrosus se spécialise dans la chaîne d'approvisionnement alimentaire et pharmaceutique et travaille à garantir la qualité et la sécurité des produits tout le long de la chaîne. Pour chaque cas, nous avons exposé le système de suivi et étudié la manière dont la blockchain a été intégrée dans leur écosystème. Ces deux startups ont développé un système qui combine l'internet des objets (IoT), la technologie blockchain et des capteurs en temps réel pour assurer la traçabilité et contrôler le flux des produits tout au long du processus de distribution. Leur objectif est d'optimiser la visibilité de la chaîne d'approvisionnement et maintenir la

qualité des produits. En introduisant la blockchain dans leur écosystème, ces startups visent à conserver la sécurité et l'authenticité des produits et prévenir la contrefaçon. En combinant les résultats théoriques avec ceux tirés de cas réels, nous avons pu établir une liste de critères pour guider les acteurs de la chaîne d'approvisionnement dans la mise en place d'un système de traçabilité de bout en bout dans un système de chaîne d'approvisionnement basé sur la blockchain, afin d'améliorer la transparence de la chaîne et garantir que seuls les produits authentiques et de hautes qualités atteignent le consommateur. Au niveau du système de suivi nous devons :

- Sélectionner les étiquettes et dispositifs de suivi adaptés au produit (Il est souvent nécessaire d'utiliser plusieurs dispositifs de suivi pour répondre à toutes les exigences de conformité).
- Choisir la technologie de communication sans fil (RFID, BLE, NFC, GPRS ou 3G) adaptée aux cas d'utilisations dans un environnement de chaîne d'approvisionnement. Plusieurs facteurs doivent être pris en considération, notamment : la portée de communication, le débit de données, le coût, la consommation d'énergie et la sécurité.
- Répondre aux failles de sécurité associées aux technologies de communication sans fil à travers l'implémentation de techniques cryptographiques, notamment le chiffrement et la signature des données collectées et échangées.
- Authentifier les étiquettes et les dispositifs de suivi afin de créer un système de traçabilité fiable.
- Adopter des dispositifs de suivi avec fonction d'enregistrement hors ligne, où les données sont stockées localement jusqu'à ce qu'elles puissent être téléchargées sur la blockchain.

Après la mise en place d'un système de suivi efficient, nous devons étudier comment l'intégration de la blockchain dans l'écosystème de la chaîne d'approvisionnement aide à relever le défi des contrefaçons. Les études de cas d'Ambrosus et de Modum montrent que la blockchain a été intégrée au sein de leurs architectures afin d'améliorer le système de suivi et la gestion des données. Ainsi la blockchain permet de :

- Assurer la transparence, fiabilité et l'intégrité des données, relatives aux produits, collectées tout le long de la chaîne d'approvisionnent. Ceci réduit les risques d'activité frauduleuse et permet une prise de décision proactive.

- Assurer l'authenticité des dispositifs de suivi.

- Etablir un niveau de confiance plus élevé entre le producteur et le consommateur, grâce à une visibilité plus large qui aide le consommateur à valider l'authenticité du produit acquis.

- Assurer l'accessibilité des données à toutes les parties prenantes de la chaîne.

- Assurer la traçabilité complète du produit ce qui facilite le rappel de produits en cas de malfaçon ou de qualité sous standards.

- Faciliter les audits grâce à la possibilité de surveiller et vérifier les différentes transactions enregistrées.

Cependant, malgré tous les avantages qu'offre la blockchain, celle-ci présente des limites qu'on doit prendre en compte avant de la déployer dans la chaîne d'approvisionnement.

- Limites en matière de stockage : nous devons adopter une architecture de stockage hybride pour gérer les volumes importants de données collectées afin d'éviter de dégrader les performances de la blockchain.

- Limites en matière de performance : nous devons sélectionner la blockchain en fonctions de différents critères clés, notamment le débit (transactions/seconde), la latence (seconde), l'évolutivité et la consommation énergétique. La performance de la blockchain est influencée par la taille des blocs et l'algorithme de consensus adopté.

- Limites en matière de sécurité : la sécurité est liée à la robustesse du consensus et le type de blockchain. Par exemple, une blockchain qui utilise la preuve de travail (PoW) ou la preuve d'enjeu (PoS) comme algorithme de consensus est vulnérable à l'attaque des 51%.

S2- Défis de la Chaîne de Prestation des Soins de Santé : Parcours du Patient

Dans le contexte actuel des soins de santé, les patients peuvent se rendre dans différents établissements médicaux pour des consultations ou traitements. Par conséquent, les données du patient sont fragmentées et cloisonnées ce qui :

- Obstrue la traçabilité des soins dispensés au patient;
- Entrave le partage des données médicales entre les différents acteurs pouvant intervenir auprès du patient;
- Compromet la confidentialité des informations personnelles des patients.

Afin d'aborder les défis de la fragmentation des données, nous avons adopté l'étude de cas comme stratégie pour identifier des solutions potentielles et évaluer leur efficacité dans un contexte réel. Nous avons ainsi examiné l'évènement indésirable médicamenteux, qui représente un défi du système de santé libanais en raison de la fragmentation des soins et de la coopération inefficace entre les différentes entités du système de santé [21], [22]. Les prescriptions électroniques ont été introduites dans le but d'améliorer la sécurité des patients en réduisant les risques d'erreurs de transcription ou d'interprétation. Elles visent également à réduire les contraintes liées à l'iatrogénie médicamenteuse tel que les : interactions médicamenteuses, allergies médicamenteuses [23], [24]. Cependant, plusieurs contraintes ont été soulevées [25], [26] :

- En matière de sécurité, la plupart des systèmes de prescription électroniques adoptés sont centralisés, ce qui compromet la vie privée des patients et expose leurs données à des menaces potentielles.
- En matière d'architecture du système, tous les systèmes de prescription électroniques ne présentent pas les mêmes fonctionnalités. Certaines fonctionnalités jugées importantes ne sont pas toujours disponibles, notamment l'historique des prescriptions du patient, les allergies du patient, un système d'aide à la décision clinique (CDSS) pour prévenir les allergies et les interactions médicamenteuses.

En raison de l'intérêt croissant pour l'intégration de la blockchain dans le secteur de la santé, plusieurs

propositions ont émergé intégrant la blockchain comme outil pour résoudre certains défis liés à la gestion des médicaments délivrés au patient (tel que : surveillance des opioïdes, gestion des allergies des patients). Toutefois, ces solutions existantes sont limitées en matière de portées et de fonctionnalités et ne sont pas conformes aux règles de confidentialité de la Health Insurance Portability and Accountability Act (HIPAA). L'objectif de ce travail est d'améliorer la gestion des données des patients, notamment leurs prescriptions médicamenteuses, afin d'éviter les interactions et allergies médicamenteuses à la prise du médicament proposé. Nous devons mettre en place un environnement d'échange de données de santé qui assure la fiabilité, la transparence et la sécurité des données échangées tout en étant conforme à l'HIPAA. Selon HIPAA, les patients ont des droits sur leurs données, on cite [27]:

- Le droit droit d'accès à leurs données médicales personnelles;
- Le droit de demander la correction de leurs données médicales personnelles;
- Le droit de demander un récapitulatif des divulgations de leurs données médicales personnelles;
- Le droit de limiter certaines utilisations et divulgations de leurs données médicales personnelles.

Pour atteindre cet objectif, nous avons abordé deux questions de recherche :

- Dans quelle mesure la technologie blockchain répond-elle aux limites associées aux prescriptions électroniques ?
- Comment l'intégration synergique des technologies du Web sémantiques et de la blockchain crée un cadre sécurisé de gestion des prescriptions médicamenteuses et des allergies des patients, présentant ainsi une approche alternative aux systèmes traditionnels de prescriptions électroniques ?

Afin de réduire les risques liés aux prescriptions médicamenteuses notamment les allergies et interactions médicamenteuses, nous avons proposé un modèle conceptuel d'un système de prescription qui combine la technologie blockchain aux technologies du Web sémantique. Nous avons utilisé la blockchain pour stocker et partager les prescriptions des patients ainsi que leurs allergies médicamenteuses avec les médecins concernés. La blockchain offre la disponibilité, l'intégrité, l'immuabilité et la transparence des données partagées [28]. Cependant, la technologie blockchain ne suffit pas à résoudre tous les défis des systèmes de soins de santé, notamment en matière d'interopérabilité et de description de concepts dans un domaine

de connaissance. Nous avons alors intégré les technologies du Web sémantique dans notre écosystème afin d'assurer l'interopérabilité sémantique et de permettre l'inférence de connaissance et la prise de décision automatisée. Nous avons défini une ontologie médicamenteuse qui comprend les classes de médicaments, d'ingrédients pharmaceutiques actifs, des allergies médicamenteuses et les relations entre elles (**HasInteractionEffect**, **CauseAllergicReaction**, **HasActiveSusbtance**) nécessaires pour vérifier les affirmations suivantes :

- Le patient n'est pas allergique à l'ingrédient actif présent dans la composition du médicament nouvellement prescrit.
- Le médicament nouvellement prescrit n'interfère pas avec les médicaments pris par le patient.
- Le médicament nouvellement prescrit n'est similaire à aucun des médicaments pris par le patient.

À la suite de l'identification et de la modélisation des concepts pertinents dans notre application, nous proposons de stocker l'ontologie médicamenteuse, hors chaîne, dans une base de données orientée graphe (GraphDB) afin de pouvoir exploiter les relations entre les données ainsi représentées. Nous utilisons le contrat intelligent (Chaincode) pour nous connecter à la GraphDB [29] et interroger les données stockées suivant le modèle OWL. Les requêtes SPARQL sont utilisées pour récupérer toutes les informations nécessaires afin de notifier le médecin d'un potentiel effet indésirable, tel qu'une réaction allergique ou une interaction médicamenteuse possible avec le médicament prescrit. Après avoir vérifié l'innocuité du médicament vis-à-vis du patient, la transaction effectuée par le médecin est considérée valide et sera ajoutée à la blockchain. Nous avons choisi d'adopter Hyperledger Fabric, une blockchain privée (permissioned) où une autorité décentralisée ou centralisée régule la participation au réseau et l'accès aux détails des transactions [30]. Outre l'intégration des alertes d'allergie et d'interaction médicamenteuse pour prévenir les accidents médicamenteux, le système proposé garantit la protection de la vie privée des patients et est conforme à la loi HIPAA [27]. Différents contrôles ont été mis en place afin de maintenir la sécurité et la confidentialité des données échangées. Nous citons :

- L'implémentation de politiques d'accès dans les contrats intelligents, basée sur l'identité du participant et l'unité organisationnelle à laquelle il est rattaché, pour gérer l'accès aux ressources. Toutefois, le contrôle d'accès est également assuré par les fournisseurs de services aux membres (MSP), qui

gèrent les identités des utilisateurs et veillent à ce que seules les identités d'utilisateurs authentifiées et autorisées puissent effectuer des transactions sur la blockchain.

- L'adoption des collections de données privées afin de permettre aux patients de contrôler l'accès à leurs données et de partager ainsi leurs informations privées avec le médecin de leur choix. Cette fonctionnalité du réseau Hyperledger fabrique augmente la confidentialité des données en permettant le partage de donnée privée avec un sous-ensemble de participants autorisés sur un canal.
- Chiffrement des données stockées dans le registre (ledger) ou dans la collection de données privées.

Même si le framework proposé améliore la gestion des prescriptions médicamenteuses, de nombreuses améliorations sont encore requises afin d'assurer un meilleur suivi médical sécurisé et confidentiel. Nous devons commencer par choisir une valeur appropriée du paramètre blockToLive de façon à donner au médecin suffisamment de temps pour examiner les données partagées par le patient avant qu'elles ne soient purgées. En outre, les allergies des patients ne se limitent pas qu'aux ingrédients actifs des médicaments, mais incluent également les excipients ou tout autre substance ajoutée aux médicaments et vaccins. Nous devons donc améliorer notre ontologie pour prendre en compte les excipients et, plus largement, les vaccins. Dans cette proposition nous renforçons et sécurisons les échanges entre le patient et le médecin mais pour compléter notre solution nous devons élargir le réseau proposé en intégrant les pharmaciens et les prestataires d'assurances.

S3- Interaction entre la Chaîne d'Approvisionnement et la Chaîne de Prestation des Soins de Santé : Libérer le Potentiel des Données

Les données jouent un rôle crucial dans les chaînes d'approvisionnement et de prestation des soins de santé. Elles constituent le catalyseur des innovations notamment lorsque les professionnels de la santé, les chercheurs et les acteurs de l'industrie pharmaceutique travaillent en synergie pour révolutionner les pratiques médicales. Les laboratoires pharmaceutiques doivent avoir une bonne compréhension des besoins des patients afin de fournir les produits convenables. De même les fournisseurs de soins de santé doivent partager avec les laboratoires les informations jugées essentielles au développement de nouveaux

médicaments. De cette interaction émerge le défi de protéger la vie privée des patients. Ce défi est le résultat d'une mauvaise gouvernance des données dans les établissements de santé. En fait, les patients ignorent comment leurs données de santé sont utilisées et partagées. Les fournisseurs de logiciels de gestion de dossiers médicaux sont accusés de commercialiser les données médicales à des entreprises pharmaceutiques [31]–[34]. De plus, certains établissements de santé couverts par la loi HIPAA sont accusés de partager, avec des tiers, des données de santé identifiables [31], [35]. Sans oublier que de nombreux patients ont souvent du mal à accéder à leur dossier médical, ce qui vient à l'encontre de la loi HIPAA [36]. D'une autre part, les établissements de santé sont devenus la cible principale de cyberattaques. Maintenir la confidentialité des données constitue l'un de leurs défis majeurs, car toute violation de données porte atteinte à la dignité et sécurité des patients [37]. Les attaques par ransomware et les accès non autorisés sont les principales causes des violations de données dans le secteur de la santé. La gouvernance des données est alors nécessaire pour atténuer ces risques.

La technologie Blockchain s'est imposée comme une solution pour garantir le partage sécurisé des dossiers de santé, en remédiant aux vulnérabilités inhérentes de la centralisation des données de santé [38]–[40]. Elle est utilisée comme plateforme de partage et de gestion des droits d'accès aux dossiers de santé tout en garantissant un historique médical complet du patient [28]. Bien que la technologie blockchain présente de nombreux avantages pour les applications de soins de santé, plusieurs défis restent à relever, notamment les défis en matière de confidentialité des données dans un écosystème régit par la blockchain. Parmi les différents défis qui se posent on cite :

- Le caractère immuable de la blockchain qui va à l'encontre du droit à l'oubli du Règlement Général sur la Protection des Données (GDPR).
- Le défi de garantir la protection de la vie privée des données enregistrées sur les blockchains publiques et privées. D'une part les blockchains publiques exposent les données à tous les utilisateurs du réseau. D'une autre part les blockchains privées créent un réseau réservé aux membres où chaque participant peut vérifier les transactions ainsi que les identités sans nécessiter d'intermédiaire tiers. Ainsi le vrai défi revient à gérer les accès de contrôle de façon à limiter le niveau d'information accessible afin de maintenir la confidentialité et l'anonymat. De même, la conciliation entre la gestion de l'identité et l'anonymat constitue un défi majeur notamment pour les blockchains privées.
- Le caractère sécurisé de la blockchain repose sur des techniques cryptographiques, ce qui entraine le défi de la gestion et la sécurisation des clés cryptographiques. La compromission des clés privées

induit des risques d'utilisation frauduleuse.

Par conséquent, comment pouvons-nous aborder les limites d'un système basé sur la blockchain en matière de confidentialité et de préservation de la vie privée ?
L'objectif de ce travail est de répondre aux divers défis exposés en proposant un système de partage transparent des données de santé entre les différents acteurs du système, notamment les patients, médecins et chercheurs dans l'industrie pharmaceutique. Le système proposé doit garantir la protection des données des patients contre des accès illicites ou non-autorisés et à empêcher tout usage de donné sans le consentement explicite des patients. Le système envisagé doit aussi garantir la confidentialité et la sécurité des données de santé des patients tout au long de leur cycle de vie.
Afin de surmonter les limites du système actuel, nous proposons une nouvelle approche de la gouvernance des données à travers le développement d'un framework basé sur la blockchain Hyperledger Fabric pour libérer le potentiel des données tout en minimisant le risque de violation et d'accès non autorisé aux données. La blockchain Hyperledger Fabric est un réseau privé où l'ensembles des utilisateurs et des composants ont des identités connues. Il apporte aux membres d'un même réseau la flexibilité et la sécurité nécessaire pour limiter l'accès aux transactions uniquement aux participants appropriés possédant les bonnes clés de cryptage. Cependant, l'environnement de confiance créé par la blockchain Fabric n'est pas suffisant à résoudre les multiples violations des données portant atteinte à la vie privée des patients. Ainsi, on propose l'intégration des concepts de classification et de ségrégation des données, conjointement aux techniques cryptographiques et non cryptographiques pour renforcer les pratiques de gouvernance des données en matière de sécurité et de confidentialité. Dans ce travail, nous nous sommes intéressées au dossier médical qui recueille essentiellement les informations administratives et les informations médicales.

Classification des données : nous avons implémenté la classification des données pour personnaliser les mesures de sécurité afin de garantir que chaque participant n'a accès qu'aux données essentielles à l'accomplissement de son activité. Nous avons classé les informations du dossier médical en fonction de leur pertinence en vue de la réalisation d'une activité médicale, notamment : activité de soins des patients, activité de soins d'urgence et activité de recherche clinique ou une activité administrative dans un établissement de santé. Lors de la classification des données et conformément aux directives de désidentification HIPAA, nous avons tenu compte de la méthode Sphère de sécurité (Safe Harbour), qui consiste à supprimer 18 identifiants spécifiques de l'ensemble des données, dans le but de réduire le risque

d'identification des patients. Les informations ont été classées en quatre catégories : critique, obligatoire, facultative et restreinte.

Ségrégation des données : nous avons implémenté la ségrégation des données afin de permettre un contrôle d'accès plus granulaire aux données stockées et partagées, empêchant ainsi qu'un seul acteur de l'écosystème puisse contrôler l'ensemble des dossiers des patients. Nous avons non seulement répartie la gestion des données de santé entre les différents acteurs du système mais également fragmenter les données classifiées des patients en des lieux de stockage distincts. Ainsi, les données sont réparties en fonction de leurs classifications entre ces trois emplacements de stockages : le registre, le système de fichier interplanétaire (IPFS) et la collection de données privées. Deux IPFS privés ont été déployés, stockant respectivement les informations administratives des patients et leurs informations cliniques désidentifiées. L'utilisation de plusieurs emplacements de stockage a permis une meilleure gestion des accès, permettant un contrôle plus fin et minimisant ainsi le risque de compromission des données. Par ailleurs, nous avons séparé les activités de santé en implémentant deux canaux dans notre réseau Fabric. Le canal hospitalier qui héberge les échanges liés aux activités de soins des patients et de soins d'urgence. Le canal de recherche et développement pharmaceutique (R&D pharmaceutique) qui héberge les échanges liés aux activités de recherche clinique. En déployant deux canaux, on a créé deux sous-ensembles de membres réseaux ou les membres de chaque canal peuvent effectuer des transactions en mode privé. Les canaux fournissent l'isolement de données et la confidentialité vu que chaque canal a ses propres politiques d'accès qui réglementent l'accès à ses ressources notamment les transactions, les contrats intelligents, l'état en cours du registre appelé world state. La confidentialité est aussi renforcée à travers l'intégration des collections de données privées qui permette d'isoler davantage des données spécifiques du reste des membres du canal. Ces données sont stockées dans une base de données privées accessible que par les participants ayant la permission de les visualiser au sein d'un même canal. Seul le hachage des données est enregistré dans le registre du canal. Le hachage permet de vérifier l'intégrité des données privées.

Désidentification des données : nous avons introduit la désidentification des données dans notre solution pour protéger la vie privée des patients. En se basant sur la méthode Sphère de sécurité, nous avons dissocié les données personnelles, notamment les éléments permettant d'identifier le patient, des données médicales. Les données désidentifiées sont stockées sur IPFS, tandis que les données personnelles sont reparties entre des collections de données privées distinctes. Etant donné que chaque entité sur le

réseau doit être identifiée pour effectuer des transactions, nous avons attribué à chaque participant un identifiant à partir duquel une identité et un certificat sera généré. Le choix de l'identifiant dépend du rôle du participant et de l'organisation à laquelle il est affilié. Par conséquent, il est impossible, à un utilisateur non autorisé, d'identifier l'identité réelle du patient en suivant leurs interactions sur le réseau. De même, nous avons assigné à chaque patient plusieurs identifiants répartis à travers les différents emplacements de stockage. Ainsi, chaque sous-ensemble de données est accessible par l'un de ces identifiants. Cette approche obscurcit l'association entre les différents fragments de données. Il est de même difficile à un individu non autorisé de reconstituer le dossier complet d'un patient à partir d'une seule source.

Contrôle d'accès : nous avons développé des contrats intelligents qui appliquent des politiques d'accès basées sur les attributs spécifiques des participants au réseau, tels que le rôle de l'utilisateur, l'affiliation de l'organisation et le type de transaction. De plus, nous avons donné aux patients le contrôle sur les autorisations d'accès, partageant leurs données avec le médecin de leur choix. De même, c'est libre à eux de participer à des essais cliniques avec la liberté de partager leurs données désidentifiées avec les laboratoires de recherche.

Mise à part l'implémentation des concepts de classification et ségrégation des données ainsi que l'adoption de méthode de désidentification et la mise en œuvre de contrôle d'accès, il est impératif de chiffrer les données de santé toujours dans le but de maintenir la confidentialité. Afin de sécuriser les données du patient, nous avons utilisé plusieurs clés symétriques. L'utilisation d'une seule clé symétrique pour accéder aux données du patient pourrait compromettre la confidentialité et augmenter le risque d'accès non autorisé. Ainsi, le framework proposé, à base de la blockchain Hyperledger Fabric, facilite l'échange de données de santé entre les différents acteurs du système de soins de santé tout en respectant les règles de confidentialité HIPAA et le droit à l'oubli GDPR. Cependant, notre solution présente de nombreux défis, notamment :

- La gestion des clés symétriques introduites dans notre système ;
- Le coût du déploiement ;
- La sécurité des fonctions de hachage qui pourraient être exploitées en cas de disponibilité d'ordinateurs quantiques puissants. Les données hachées se trouvant sur le registre seront alors compromises. Il serait plus sûr d'introduire un processus de randomisation avant d'ajouter la valeur de hachage sur

la blockchain.

Conclusion

Bien que ce travail ait fourni une base solide pour débloquer le potentiel de la technologie blockchain, relevant ainsi certains de ses défis dans le secteur des soins de santé, d'autres défis n'ont pas encore été abordés. Plusieurs travaux futurs peuvent être envisagés, entre autres :

1. Améliorer l'interopérabilité inter-blockchain. Le secteur de soins de santé est composé de multitudes acteurs qui interagissent les uns avec les autres, il est ainsi peu probable qu'ils adoptent tous la même plateforme blockchain. L'écosystème de santé est alors constitué de différents réseaux de blockchain chacun caractérisé par sa propre architecture, mécanisme de consensus, niveau de permission et langages de programmation. L'adoption de technologie cross-chain (tel que : swaps atomiques, Sidechains, les ponts transversaux) est essentielle pour favoriser la collaboration et l'interopérabilité entre ces différents réseaux. Cependant, le défi consiste à concevoir un système interopérable qui garantit à la fois la sécurité et la gouvernance. Ainsi, les travaux futurs devront se concentrer sur ces défis, afin de rendre l'écosystème décentralisé plus résilient et adaptable.
2. Monétiser les données de santé, approche pour inciter les patients à participer à la recherche clinique. La vente des données de santé représente un marché lucratif, où les patients en contrôles de leurs propres donnés pourraient en bénéficier en le partageant volontairement avec les chercheurs cliniques. Cependant, la monétisation des données de santé soulève plusieurs enjeux juridiques et éthiques qui nécessitent une attention particulière et peuvent entraver l'adoption d'une telle approche. Ainsi, les travaux futurs consistent à évaluer ces enjeux et à examiner la viabilité d'une telle proposition, permettant au patient de bénéficier des avantages économiques tout en respectant ces droits et sa vie privée.
3. Améliorer la protection de la vie privée des patients dans le cadre de l'intégration de l'intelligence artificielle dans les technologies médicales portables. L'adoption de ces technologies a révolutionné les soins de santé, offrant aux patients des services personnalisés, mais au détriment de la protection de leur vie privée. Dans cette thèse, nous nous sommes concentrés sur les défis liés à la gestion

des dossiers médicaux. Bien que ces données soient sensibles et de grande valeur, le cycle de ces données ainsi que le rôle des principaux acteurs vis-à-vis de ces données étaient bien définis, ce qui nous a permis de contenir les risques liés aux données. En revanche, lorsque nous nous aventurons dans le domaine des technologies portables et de l'intelligence artificielle, le dynamisme change considérablement et le cycle des données ne peut pas être facilement prédit. Ainsi, les travaux futurs, consistent à relever le défi de trouver le bon équilibre entre le respect de la vie privée et l'usage des données en tant que monnaie pour améliorer l'expérience de l'utilisateur.

Chapter 1

Introduction

1.1 Background and motivation

In today's healthcare ecosystem, the patient is at the center of a complex network in which every interaction among the different stakeholders contains opportunities for error. Information may be collected, recorded, or communicated inaccurately. Technical failures, such as power outages and system crashes, can impact the delivery of healthcare services or goods and increase the risk of error. Additionally, security breaches can significantly harm the trust relationship among the different stakeholders of the healthcare system. Much like a traditional supply chain, the healthcare ecosystem is a network of organizations, people, activities, information, and resources involved in providing and delivering a product or service from supplier to customer. The healthcare supply chain has grown so enormous and complicated that it is hard for anyone, let alone a single company or organization, to comprehend all its processes [2]. It involves multiple independent stakeholders working to deliver high-quality goods and services to healthcare providers and patients. The main players in the healthcare supply chain are [1]:

- Health providers: They include health professionals (physicians, nurses, etc.) and health facility organizations licensed to provide healthcare diagnosis or treatment (hospitals, clinics, medical laboratories, etc.)

- Vendors: They include pharmaceutical manufacturers, medical equipment suppliers, and healthcare software providers.
- Payers: They include the government and/or the insurance companies, which are both involved in financing healthcare maintenance and restoration.
- Research and development laboratories: They include academic, industry, and government facilities, which are involved in the development of new medical treatments and technologies.

As highlighted by Kitsiou S. *et al.*, the healthcare supply chain is very fragmented, complex , diverse, and dynamic [2]. It consists of chains of operations that must be performed by a single provider or in collaboration among different providers to produce a particular service or product [3]. Hence, the chain of operations is a process with multiple customers but one ultimate consumer: the patient [3]. Healthcare supply chain management is fundamental for maximizing patient outcomes, improving stakeholder coordination, controlling costs, and reducing human error. Since healthcare delivery can also be viewed as a supply chain, it is crucial to highlight that "supply chain" will have a different connotation when used to describe the flow of goods versus the flow of patients in the healthcare system [3]. In the present thesis, we identify the latter as the healthcare delivery chain or care chain to avoid ambiguity. We make the following distinction:

- Healthcare supply chain management: It focuses on the coordination and management of activities related to the movement of goods or products from manufacturers to retailers and consumers [2], [3], [9]. It includes managing the flow of medical supplies and equipment, pharmaceuticals, and other healthcare products through the different supply chain stages. It also involves managing the flow of information and data related to these activities, such as inventory levels, quality control data, orders, and tracking information [2], [3], [9].
- Healthcare delivery chain management: It focuses on managing and tracking the movement of patients through the healthcare system, from the moment they seek medical care until they are discharged [1], [3]. It includes appointments coordination, treatment planning, diagnosis, and follow-up care. It also involves managing the clinical workflow, which describes how data moves through the health information systems and to whom. The latter include patient data, medical histories, medications, and communication among healthcare providers such as doctors, nurses, and other medical professionals

[1], [3].

In addition, each of these chains is characterized by a third flow: the financial flow. The financial flow encompasses the different financial operations among the various stakeholders in the healthcare system. It includes the funds from public or private payers and patients [2], [3].

Even though the healthcare delivery chain is dissociated from the healthcare supply chain at the level of the chain of operations, they are both linked through the patient. Medical manufacturers must have a good understanding of the needs of both patients and healthcare providers to provide efficient and relevant products. In addition, healthcare providers must collaborate with manufacturers by sharing all information deemed essential for the sole purpose of helping them develop new products to improve patient outcomes [3]. Thus, a massive amount of data is generated daily in the healthcare sector, both structured and unstructured. These data encompass data generated by smart devices which can track goods throughout their entire lifecycle to ensure they meet strict quality and safety standards or assist in managing various health needs, ranging from remote health monitoring to symptom tracking and predictive healthcare. Therefore, there is a need to enhance information management and system integration to maintain trust among the different health stakeholders of the supply and care chains [2]. Such a goal is achieved through a policy that emphasizes chain transparency, ensuring the availability of accurate and relevant information in the correct format to the proper users at the right time to support timely decision-making across the entire chain [1], [2]. Data governance is essential for producing information and knowledge that can improve patient outcomes and support innovation in the healthcare sector. It requires a thorough understanding of the data-related risks involved in the management process [41]. The principal causes of potential data risks in the healthcare industry are centralized management systems, insider threats (notably employees, unsecured third-party vendors, or partners), disconnected data, and poor data security, which can result in data breaches, privacy violations, and cyberattacks [1], [4]–[6].

As revealed by Stoshi Nakamoto, blockchain technology emerged in 2008 to serve as the shared ledger of the cryptocurrency Bitcoin [7]. Unlike traditional currencies, Bitcoin eliminated the need for intermediaries and provided an efficient way to record transactions' information [42]. What the blockchain brought to financial services is security, immutability, transparency, and the ability to excise the middleman [8]. Used to record any transaction and to track the movement of any asset, the blockchain revolutionized

the traditional business network. Many sectors - notably in healthcare, insurance, government, supply chain management, and Internet of things - are likely to be transformed by the blockchain [13]. The main goal of this thesis is to study the integration of blockchain technology into the healthcare supply and care chains and its role in improving data transparency, promoting accountability of healthcare actors' activities, and preventing security breaches that compromise private or personal data.

1.2 Thesis overview and organization

Given that healthcare is a unique industry with its own set of regulations and requirements, it is imperative to consider the multiple perspectives and priorities behind each interaction among its various stakeholders. This consideration extends to the type of data handled when implementing data privacy and security measures [1]. In fact, the supply and care chains have unique challenges, processes, and priorities, making a one-size-fits-all approach ineffective. A thorough assessment of each chain is necessary to determine the specific challenges that must be addressed and improved. By adopting a tailored approach, we can identify how blockchain technology can be introduced into each of the supply and care chains to improve their efficiency by ensuring the confidentiality, integrity, and availability of sensitive information and protecting the privacy rights of patients and supply chain partners.

Several studies have proposed the blockchain as a potential solution to alleviate some issues associated with healthcare scenarios since the blockchain provides an untampered/unalterable record of transactions [28]. However, introducing the blockchain into the healthcare ecosystem has also brought some challenges that should be addressed, notably in terms of scalability, storage, security, and privacy. To what extent a blockchain-based solution can handle the huge amount of healthcare data while being compliant with the Health Insurance Portability and Accountability (HIPAA)? And at what cost?

In this thesis we will focus on a three healthcare ecosystem challenges that compromises the safety and/or privacy of patients:

Challenge 1. Healthcare Supply Chain: Tracing the Drug Journey

Challenge 2. Healthcare Delivery Chain: Tracing the Patient Journey

Challenge 3. Supply Chain Meets Care Chain: Navigating Healthcare Privacy Challenges

The thesis is divided into four main chapters. In **Chapter 2**, we present a rich literature review that identifies and assesses the role of the blockchain technology in both the supply and care chains. In this chapter, we answer the following questions:

- What is a blockchain?
- How does blockchain improve the supply chain management?
- How does blockchain improve the healthcare delivery chain?

In **Chapter 3**, we address the centralized supply chain management system that exposes the supply chain to corruption, fraud, and tampering. This study describes how the blockchain, with its decentralized architecture, can be integrated into the supply chain architecture to create a reliable, transparent, authentic, and secure system. To reach this goal, we study the benefits of introducing the blockchain to the supply chain and the challenges encountered in a blockchain-based supply chain management ecosystem. We combine theoretical and real-world application studies to build our theory about the requirements for an efficient blockchain-based supply chain.

In **Chapter 4**, we address health data fragmentation issues by adopting a case study approach to examine the adverse drug interactions/reactions challenge in the Lebanese healthcare system. We propose a conceptual framework that enhances the management of patients' data, notably their drug prescriptions and drug allergies, while maintaining their privacy as per HIPAA requirements. Although the initial target of our contribution is the healthcare system in Lebanon, our solution can be applied to improve any healthcare system. It combines the usage of a blockchain with Semantic Web technologies, where the blockchain provides a secure infrastructure for data sharing and prevents data tampering through a tamper-proof audit log. As for the semantic description of data, it brings a common, shared understanding of the data at hand, boosting knowledge discovery and automating decision making. In addition, with the Semantic Web technologies introduced alongside the blockchain to enhance the management of the healthcare delivery chain, we have found it essential to conduct a thorough literature review of the role of

the Semantic Web in addressing healthcare challenges.

The studies conducted in **Chapter 3** and **Chapter 4** show the role of blockchain technology in maintaining the availability, traceability, and integrity of collected data, whether linked to a patient or a pharmaceutical product. Although the blockchain revolutionizes the healthcare industry, a blockchain-based ecosystem still faces multiple challenges, notably privacy issues, not to mention the limited storage capacity of the blockchain. This requires the integration of off-chain storage to handle massive data, enhance blockchain performance, and address the challenge introduced by the "right to be forgotten," a fundamental right proclaimed under Articles 17 and 19 of the GDPR. Hence, the blockchain-based healthcare ecosystem faces new challenges, notably scalability, storage, security, and privacy. By the latter, we mean the challenge of maintaining the privacy and confidentiality of the data exchanged among different healthcare actors stored on-chain or off-chain.

Thus, in **Chapter 5**, we address the storage limitation of the blockchain in addition to the confidentiality and privacy concerns associated with the blockchain-based health information exchange systems. Data confidentiality is one of the biggest challenges faced by healthcare organizations since a data breach can affect a person's dignity, with severe consequences for the organization itself. Data is the lifeblood of the healthcare supply and care chains. It pulses through every aspect of the industry, from healthcare logistics to patient care. It is the key to innovation, especially when the pharmaceutical supply chain meets with the care chain for transformative research journeys. Patients are unaware of the volume of information recorded about them, these data's value in clinical studies, and the way they are shared among healthcare actors. Besides, when patients need to access or get their medical records, healthcare providers meet their requests with delays and fees while being selective about the shared information. Health record retention by health providers will keep them in control of information with high commercial and research value. Data governance is necessary to mitigate these risks, especially when dealing with sensitive data such as health data. We present in **Chapter 5**, therefore, a novel approach to data governance through the development of a blockchain-based framework that aims to unlock data's potential while minimizing the risk of data breaches and unauthorized access. The framework involves three critical components: data classification, data segregation, and data access control, to ensure proper management, protection and utilization of data within an organization. Even though our proposed framework is applied within the healthcare ecosystem, we present a clear roadmap to facilitate its adoption across diverse domains. Our approach demonstrates

how data accessibility can be enhanced while safeguarding data privacy and upholding its availability, integrity, and confidentiality.

Finally, **Chapter 6** concludes the thesis with a summary of the research outcomes, the strengths and limitations of the work, and the perspectives for future research.

Chapter 2

Blockchain Technology

2.1 What is a blockchain?

A blockchain is a distributed ledger that records and shares all transactions that occur within the blockchain network. The blockchain network consists of multiple nodes that maintain a set of shared state and perform transactions modifying the states [43]. Transactions must be validated by the majority of network nodes, before being ordered and packaged into a timestamped block. This mining process depends on the consensus mechanism adopted by the blockchain network [28]. Before adding the new suggested block to the chain, all networks' nodes verify that the block contains valid transactions and references the correct previous block via a cryptographic pointer.

The blockchain network can be categorized either as permissionless or as permissioned network. A permissionless blockchain, is an open distributed ledger where any node can join the network and where any two peers can conduct transactions without any authentication from the central agency [28]. A permissioned blockchain is a controlled distributed ledger, where the decision making, and the validation process are handled by a centralized or decentralized authority [28]. A certificate authority determines who can join the network. All nodes are authenticated, and their identity is known to other nodes [43].

Figure 2.1 below shows the blockchain data structure. The first block is known as the genesis block. A block consists of a header and a body. The block body contains the list of transactions [28]. The number of

transactions within a block is related to the block and transaction's size. The block header contains various fields, mainly the block version indicating the set of rules which should be followed for validation , a hash of the previous block header, a timestamp, the Merkle tree root hash that represents the hash value of all the transactions in the block [28], [44]. The nonce and target are block header fields, used for the Proof of Work protocol. It's a computational process, known as mining, where miners are the nodes that calculate the block header hash. A block is accepted by all nodes if a miner finds a nonce such as: hash (block header) < difficulty target. The nonce is a 32-bit field that is incremented until the equation is solved [28], [44].

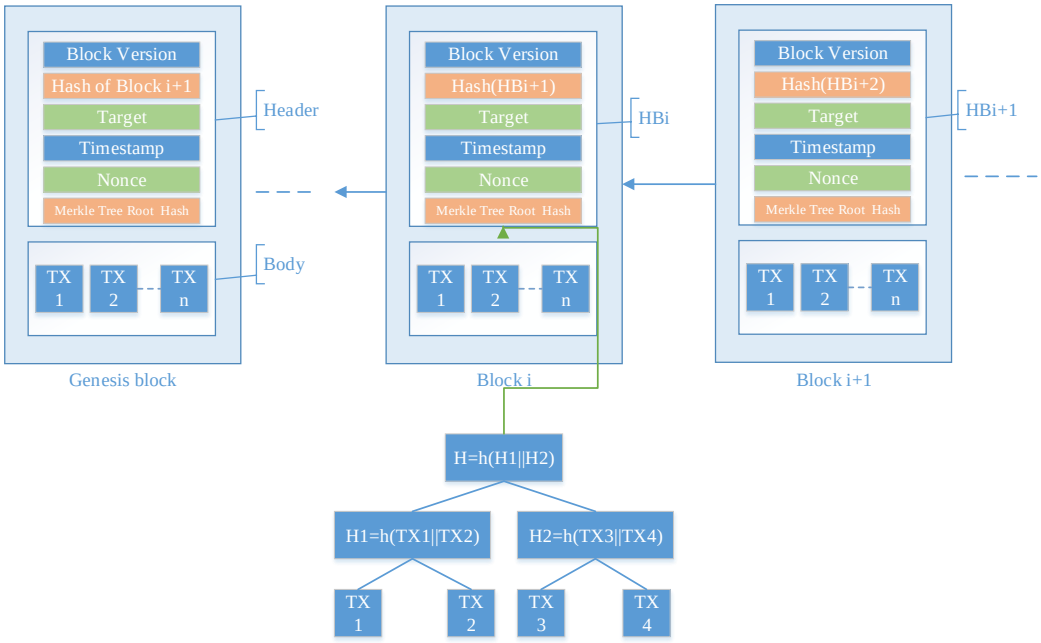


Figure 2.1: Blockchain structure

Apart from being a distributed shared ledger, blockchain is also defined by three key concepts: consensus, smart contract, and cryptography [42], [43].

1. A consensus is an agreement that helps a decentralized network to authenticate and validate a value or a transaction. It ensures that all network nodes share the same data and prevents malicious actors from manipulating the data [45]. A consensus mechanism is defined by the following parameters: integrity, authentication, non-repudiation, byzantine fault tolerance, decentralized governance, quorum structure and performance [46]. The type of consensus protocol depends on the blockchain type. For

example, Bitcoin, a public ledger, uses Proof of Work, a computational expensive mining protocol to work around the Sybil attack, where a minority can control the whole network. In a permissioned blockchain, one organization or a group of organizations determine the consensus process. A node needs to be certified to join the consensus process [28]. In that case, Proof of Work is unnecessary and is an expensive way to reach consensus because all participants are authenticated.

2. Smart contracts are self-executing scripts stored on the blockchain. When performing a transaction, smart contracts are invoked to execute the term of a contract/procedure on every node in the network [28]. Hence, every node in a blockchain network must agree on the inputs, outputs and states affected by the smart contract [43]. Satisfying common contractual conditions, such as payment terms or confidentiality, minimizes the need for trusted intermediaries [47].

3. Cryptographic techniques are used to ensure integrity, authenticity, immutability and nonrepudiation of the blockchain ledgers since even an authenticated node can act maliciously [28]. The state root hash and the hash pointers are combined to secure and track all the historical changes made to the global state [43], [44]. The purpose of the root hash of the hash tree is to detect data tampering and to validate the transaction efficiently [28]. To verify any transaction, we need to check the hash tree path related to the requested transaction. Any modification in a specific transaction will be instantly detected [43]. The purpose of the block header hash is to verify the integrity of the block and of the transactions, and to form the chain link by embedding the previous block hash in the current block header. Transactions' block cannot be modified or deleted, once appended to the blockchain. Any modification in a specific block will invalidate all subsequent blocks [28], [43]. The asymmetric cryptography is used to provide integrity, authentication and nonrepudiation into the blockchain network. A user's node must sign the transaction before broadcasting it to the network [28], [43], [44]. Each user generates a key pair. The private key is used to encrypt the hash value derived from the transactions, and the public one is used by a peer node to verify the transaction's authenticity. Note that in a permissioned blockchain, an access control layer is added. For example, in Hyperledger, arbitrary policies are implemented to control users' access to the blockchain, thus adding more security to the network [43].

2.2 How does blockchain improve the supply chain management?

The manufacturing of goods is becoming complex due to the increased number of intermediaries between the producer and the final consumer. Globalization and market expansion pushed companies to expand their products portfolios and life cycle, to meet new markets requirements. Hence, there's little knowledge of the product origins, processing or shipping journey [48], [49]. The challenge becomes not only quantitative but also qualitative. The main challenge of the supply chain remains in the traceability and data management system. The management of information system in most sectors notably in healthcare, financial, food, and education is centralized. Transactions, decision-making, and storage system are controlled by third-party intermediaries. However, a centralized management system could represent a threat to data integrity, availability, and resiliency, leaving the system subject to corruption fraud and tampering [14]. A trusted ecosystem needs to be created between the suppliers and their customers. This is achieved by a policy that focuses on the transparency of the chain to ensure product traceability, where accurate data collection and secure data storage are required.

A good traceability system aims to minimize the production and distribution of unsafe or bad quality products by improving the labeling and tracking systems. The track and trace systems have evolved from paperwork to Internet of things (IoT) hardware and sensors [50]–[53]. The principal components of a tracking system are the tag, the tracer, and the sensor. A tag is a label set on the top of a product or a package that identifies the product. Passive Radio-frequency identification (RFID) and Quick Response Code (QR code) are examples of tagging systems. A tracer is a substance introduced into a product or its natural feature, used to provide information about the course or the process that involved a product, thus certifying its quality. A sensor is a device that detects environmental changes such as light, heat, motion, moisture, pressure, etc. The detected events are then sent to other electronic devices over the network for processing. However, tracking devices are sometimes compromised and subject to cloning. An attacker can clone an RFID tag attached to a genuine product. Cloned tags on counterfeit products [15] can mislead the consumers and endanger the consumers' safety in a medical or food industry [54]. For producers, cloned tags can damage the company's reputation and cause severe economic losses in the logistics industries. Resolving clone attacks issue is achieved either through a prevention strategy based on developing either clone attack detection technique or a tag distribution schemes in order to prevent an attacker from copying

the tags' content [15]. According to Toyoda *et al.*, none of these proposed track and trace methods can guarantee that the product, with an attached tag, is genuine once it is placed in retails stores for sale; this is because these methods leverage the tag's secret information [15]. A blockchain-based product ownership management system was proposed, to transfer and prove the uniqueness of an RFID tag-attached products for the post supply chain. Counterfeits may be detected when the seller cannot prove the possession of the claimed product.

In supply chain area, storage and logistics management is considered a real challenge. Petri Helo *et al.* discuss in their paper the limit of centralized enterprise resource planning (ERP) technology in managing the supply chain and introduce a cloud-based solution [55]. In fact, ERP is a transaction management system that processes collected information and stores data in a single database. But ERP could not adapt to supply chain evolution and requirements especially in terms of transparency, flexibility, data accessibility and advanced decision making. A cloud-based NetMES system has been proposed to solve this issue [55]. The cloud technology is used as a platform to exchange, store and monitor information where a centralized virtual database replaces a centralized physical database. It has added a real-time interaction to the whole proposed system; however, the security and privacy of stored data remain an issue [55], [56].
Besides, in a centralized system, a single entity controls data. If this entity fails or shuts down abruptly, the whole system will crash and stop processing transactions [57]. The system is subject to fraud and malicious attack. It's not the case with the distributed ledger where a hacker cannot take advantage of a vulnerable point; if one node fails, the remaining nodes will not be affected. Note that a centralized system allows any user to modify a transaction in the ledger because there is no restriction on the operations [58]. In case the data administrator is bribed, the whole system could be subject to tampering and falsifying information [57]. In China, the agrifood loss ratio is up to 30% yearly mainly due to their centralized logistic system [59]. To reduce the losses during the logistics process and enhance food safety, Feng Tian proposed a decentralized traceability system based on RFID and blockchain. According to Feng Tian, enhancing the quality of the traceability system by integrating the RFID with other technologies such as wireless sensor network (WSN), Global Positioning System (GPS), etc. is not sufficient [59]. These technologies cannot guarantee the integrity of the collected and shared data with all supply chain members.

Integrated to improve the tracking system, blockchain strengthens trust, food safety assurance and information credibility. The RFID executes the tracing and monitoring to guarantee food quality and

safety. All relevant information is then uploaded on the blockchain to create a reliable, transparent and secure decentralized platform, where all supply chain actors can interact [59]. In case of an accident, emergency measures could be immediately taken to prevent the risk of hazard spreading.
Therefore, blockchain provides a permanent record of transactions, where data is maintained on multiple blocks that cannot be altered. All product and shipping details are collected through different technologies and validated before becoming a permanent record on the blockchain [48], [60].

2.3 How does blockchain improve the healthcare delivery chain?

To explore the key research priorities at introducing the blockchain into the healthcare delivery chain, we searched the IEEE Xplore research database for the exact keywords: "Blockchain" AND "Healthcare". We selected a total of 43 papers, published between 2017 and 2020, that proposed a patient-oriented framework that integrates the blockchain as a tool to address healthcare issues, and we reviewed the following questions:

- What healthcare issues do the authors address in the paper?
- How was the blockchain introduced into the healthcare system?
- What has the blockchain brought to the healthcare ecosystem?

The extracted data are reported in the below tables. We sorted the research papers according to their field of application, and we distinguished two main categories:

- Blockchain for digital healthcare records Table 2.1: it includes the papers that address the issues related to electronic medical records (EMR), electronic health records (EHR), personal health records (PHR), and electronic patient records (EPR).
- Blockchain for the smart healthcare ecosystem Table 2.2 : it includes the papers that address the issues related to pervasive healthcare applications such as: telehealth, telemedicine, long-term/short-term remote patient monitoring, intelligent emergency system, incidence detection, etc.

Table 2.1: Data extraction results from the articles related to the blockchain application in the digital healthcare record

	Blockchain for Digital healthcare record				
	Healthcare Challenges	Storage	Blockchain		
			Platform	Role	Impact on system
[61]	1-Centralized design 2-Tracking patients' data 3-Bad use of data by online administrators	**Off-chain:** private cloud stores the encrypted medical data. **On-chain:** address of the encrypted medical data.	-	To grant security in accessing patients'medical report (only used for data retrieval).	Data integrity, Data sharing.
[62]	1-Give patients control over their health data 2-Data availability	**Off-chain:** medical data kept in provider database. **On-chain:** list of administrative authorities, permission contracts.	Ethereum blockchain	To give patients control over their record by managing access right.	Access control, User authentication.
[63]	1-Centralized storage systems 2-Data privacy and security issues	**Off-chain:** IPFS stores the COVID-19 patient's diagnostic reports. **On-chain:** the content-address hash of the patient's report.	Consortium blockchain	To share the patients'report while preserving privacy and integrity, and easy searchability.	Data privacy, Data integrity, Data sharing.
[64]	1-Long process to access patient's health data 2-No secure way to share health records	**Off-chain:** medical record kept anonymously in a centralized database. **On-chain:** pseudo anonymous identifier, hash pointer to the index of the file, hash of the file.	Multichain permissioned blockchain	To allow secure access to medical records between institutions and to rapidly collect and view the patients' medical history.	Data integrity, Data sharing.
[65]	1-Interoperability issues 2-Fragmented health data 3-Give patients control over their health data 4-Sharing data among untrusted peers	**On-chain:** encrypted medical record, encrypted session keys.	Multichain permissioned blockchain	To provide a patient centric access control for EMR.	Access control, Data integrity, Accountability, Data sharing, Patient privacy preservation.
[66]	1-Centralized design 2-Interoperability issues 3-No secure way to share medical records	**Off-chain:** cloud server stores the encrypted patients' records and the symmetric keys. **On-chain:** attribute of the verified staff member, access verification/permission contracts.	Ethereum public blockchain	To regulate access permissions of patients' records while allowing the patients define access policies to the staff members of medical institutions.	Access control, User authentication.
[67]	1-Cloud-based approach issues 2-Data Privacy issues	**Off-chain:** cloud server stores the patients' record **On-chain:** smart contracts to monitor, identify, report all actions performed on data and revoke access in case of violation.	Implemented their own blockchain: "MedShare"	To permit data audit, access control, and data provenance.	Secure data auditing and trailing, Access control.
[68]	1-Centralized design 2-Give patients control over their health data	**Off-chain:** cloud server stores patients' EHR ciphertext **On-chain:** • keyword ciphertext • EHR indexes (file location) • data owners' account address • data providers' signature • access request/authorization transactions	Implemented their own consortium blockchain	To provide a secure, and rapid data sharing with easy data searchability.	Data integrity, Patient privacy protection, Access control, Identity & Data authentication, Data sharing, Data searching.
[69]	1-Interoperability issues	**Off-chain:** medical record kept in institutional database. **On-EMR-chain:** hash value of the medical record, hospital signature, patient signature, set of keywords. **On-PHD-chain:** data collected from individuals.	Implemented their own blockchain "BlocHIE" = EMR-Chain + PHD-Chain	To stores and share the collected healthcare data while preserving privacy, integrity, and easy searchability.	Interoperability, Patient privacy protection, Data sharing, Data integrity.

[70]	1-Interoperability issues 2-Long process to access patient's medical data 3-Centralized design	**Off-chain:** medical record kept in institutional database. **On-chain:** URI of health data, hash of the actual medical data, operations logging data (access activity).	Hyperledger permissioned blockchain	To record and trace the medical data sharing and data access operations with immutable transactions.	Secure data auditing and trailing, Access control, Data sharing, Data integrity.
[71]	1-Centralized design 2-No secure way to share medical records	**Off-chain:** medical record kept in institutional database. **On-chain:** address of the medical data, access verification/permission contracts.	Ethereum blockchain	To append or retrieve a medical record in a secure way while allowing the patient to control access permission.	Data integrity, Access control, Data sharing.
[72]	1-No secure way to share medical records	**Off-chain:** cloud server stores the medical record. **On-chain:** smart contract related to data entity listed for sale (includes: the data signature, the URL of the data stored off-chain, list of public keys that granted access, the price for the data access).	Public blockchain	To provide a secure data exchange.	Data integrity, Access control, Data sharing.
[73]	1-Interoperability issues 2-Data privacy issues 3-Give patients control over their health data 4-Long process to access patient's medical data	**Off-chain:** cloud-based repository stores all the encrypted data assets. **On-chain:** hash of the data asset's URI, contracts to control access.	Hyperledger permissioned blockchain	To provide an efficient and secure exchange of health data while allowing the patient to control access permission.	Data integrity, Access control, Data sharing, Auditing and request tracking.
[74]	1-Data sharing issues 2-Fragmented health data 3-Give patients control over their health data	**On-chain:** Electronic patient record access agreement, agreement list, patient's health record (using universal interpretable codes: LOINC and epSOS.).	Permissioned blockchain	To provide an efficient and secure exchange of health data while allowing the patient to control access permission.	Data integrity, Access control, Data sharing.
[75]	1-Fragmented health data 2-Data sharing issues 3-Centralized design	**Off-chain:** medical data resides on each peer's local database. **On-chain:** smart contracts to manage access permission.	Public Ethereum blockchain	To control permission for shared data through smart contracts and notify sharing peers on data update.	Data integrity, Access control, Tracking data updates and request.
[76]	1-Data sharing issues 2-Leakage of highly sensitive data 3-Centralized design 4-Interoperability issues	**Off-chain:** cloud-based repository stores all the patients' encrypted medical records and the extraction signature. **On-chain:** storage location URL, permissions contracts, and the operation logging data (access activity).	Consortium blockchain	To provide an efficient and secure exchange of health data while allowing the patient to control access permission.	Data integrity, Access control, Auditing and request tracking, Data sharing.
[77]	1-Data sharing issues	**Off-chain:** PostgreSQL, a relational database, used to store sample data. **On-chain:** smart contracts to manage authorization.	Hyperledger permissioned blockchain	To manage (grant/ revoke) the access of the database.	Access control
[78]	1-Leakage of highly sensitive data 2-Data sharing issues	**Off-chain:** medical record kept in EHR servers. **On-chain:** smart contracts to manage access permission, hashed EHR data pointing to the main data stored on EHR servers.	Hyperledger permissioned blockchain	To grant access to authorized parties to view or to update data on the EHR server while allowing the patient to control access permission.	Data integrity, Access control, Patient privacy protection.
[79]	1-Data analytics issues 2-Data privacy issue	**Off-chain:** medical record kept in EHR and depersonalized patients' data kept in cloud storage. **On-chain:** encrypted and signed data.	Exonum permissioned blockchain	To protect confidential data, monitor and record data on changes in medical records.	Data integrity, Data authentication, Monitoring and tracking data changes.
[80]	1-Give patients control over their health data 2-No secure record management of patients' record 3-No data access tracking 4-Misuse of health data in the emergency	**Off-chain:** PHR data kept in a secure cloud-based repository, and PHR URI kept in the client application. **On-chain:** smart contracts to manage authorization.	Hyperledger permissioned blockchain	To store and update data request, and to manage and record data access to the PHR, in a secure way.	Access control, Auditing, Data availability, Patient privacy protection.

[81]	1-Centralized design 2-No secure way to share data 3-Key escrow problem in an attribute-based encryption scheme	**Off-chain:** cloud server keeps PHR ciphertext set, and keyword index set. **On-chain:** • signature of the cloud server • hash value of all encrypted PHR • 9 different smart contracts functions (such as: add/remove user, data integrity verification) • ciphertext of the attribute private key • ciphertext of the searchable symmetric encryption scheme key	Ethereum public blockchain	To provide an efficient and secure exchange of patients' health data while allowing the patient to control access permission.	Data integrity, Access control, Data sharing, Secure key management.
[82]	1-No secure way to share data 2-Cloud-based approach security issues	**On-chain:** verified medical results	Private blockchain	To provide an efficient and secure exchange of the medical data.	Data integrity, Data sharing.
[83]	1-Access control issue	**Off-chain:** medical record kept in EHR databases. **On-chain:** 6 different smart contracts (such as: permission contract and owner contract (hashes of records and query link)).	Ethereum blockchain	To provide an efficient and secure exchange of the medical data.	Access control, Data sharing, Data integrity, Patient privacy protection.
[84]	1-Centralized design 2-Cloud-based approach security issues 3-Give patients control over their health data	**Off-chain:** IPFS stores encrypted medical data. **On-chain:** address of patient data, 3 different smart contracts operations (grant access, identify/validate request).	Ethereum private blockchain	To manage user access for ensuring efficient and secure EHRs sharing.	Access control, Data integrity, User authentication, Data monitoring Patient privacy protection.
[85]	1-Lack of standardized integration of the health system 2-Interoperability and security issues 3-Data privacy issue	**Off-chain:** medical record kept in EHR databases. **On-chain:** summary contract (access right) and record relationship contract (read/edit logs).	Permissioned blockchain	To track all events that happened to the data in the EHR systems while allowing the user to control access permission.	Access control, Event tracking.
[86]	1-Data sharing issues for research purposes	**Off-chain:** Web/cloud platforms store Patient healthcare data. **On-chain:** permission contract, registry contract, patient data contract (patient hashed healthcare data + URL to patient's data).	Consortium blockchain	To share healthcare data in a secure, private, and auditable way while allowing the patient to control access permission.	Data integrity, Auditing and accountability, Patient pseudonymity.
[87]	1-No secure way to share data 2-Give patients control over their health data	**Off-chain:** internal hospitals' database stores the personal information of the patient. **On-chain:** patients' medical questionnaire results.	-	To share securely medical questionnaire result data while allowing the patient to control access permission.	Data integrity, Data sharing, Access control.
[88]	1-Share and integrate medical data between different healthcare systems 2-Challenge to provide an accurate diagnosis 3-Security issues	**Off-chain:** cloud storage stores the data collected from IoT devices and the EHR/PHR repository keeps the medical records (including the lab test records). **On-chain:** patients' health events in ontology terms, URL pointer to the EHR/PHR repository, access permission contracts, biomarkers, and important events from IoT devices.	Ethereum blockchain	To share health data in a secure, private, and auditable way while allowing the patient to control access permission.	Data integrity, Data sharing, Access control, Auditing.

[89]	1-Share and manage personal health data (PHD). 2-Security & privacy issues	**Off-chain:** hot storage and cold storage used to store PHD. **On-chain:** smart contracts to manage all the protocols between the different system actors.	Hyperledger Fabric permissioned blockchain	To share healthcare data in a secure and private way while allowing the patient to control access permission.	Data sharing, Access control.
[90]	1-No secure way to share data 2-Fragmented health data	**Off-chain:** data repository stores all the medical data and a copy of the health authority blockchain. **On-chain:** • **Permissioned blockchain:** patient ID, encrypted link to medical records, hash of the medical data, consent log. • **Health authority blockchain:** patient ID, original and updated patient data, information about caregivers.	**Dual blockchain model:** Patient permissioned blockchain & Health authority blockchain (public or consortium)	To share medical data in a secure, private, and auditable way.	Data integrity, Consent traceability, Data sharing.

Table 2.2: Data extraction results from the articles related to the blockchain application in smart healthcare

	Blockchain for smart healthcare				
	Smart Healthcare challenges	Storage	Blockchain		
			Platform	Role	Impact on system
[91]	1-Data sharing while preserving reliable real-time service 2-Data access control 3-Enforce electronic protected health information (ePHI)-regulations restrictions	**Off-chain:** institutions' HIPAA-compliant databases store the digitally signed data collected from patient's home and the medical data from health institutions. **On-chain:** URL of the actual data, secure index, access control & data integrity smart contracts.	Permissioned Ethereum blockchain	To manage secure data sharing on sensitive patients' data while allowing the patient to control access permission.	Data integrity, Data validation, Identity authentication, Access control.
[92]	1-Vulnerabilities within current IoT model in terms of use of a central cloud server 2-Data access control 3-Limited computing and storage capacity within smart health devices	**Off-chain:** IPFS stores the encrypted health data. **On-Userchain:** hash of encrypted IoT data and symmetric keys (encrypt/decrypt IoT data & diagnosis) **On-Docchain:** hash of encrypted diagnosis	Implemented their own blockchain Healthchain = Userchain (public) + Docchain (consortium)	To provide a secure diagnosis with fine-grained access control of largescale health data while allowing the patient to control access permission.	Access control, Data integrity, Accountability, Data privacy.
[93]	1-Access control of the IoMT ecosystem 2-Data integrity and privacy issue	**Off-chain:** IPFS stores patients' health record. **On-chain:** smart contracts that handle registration of users/devices, provide access control and manage the logs produced by users.	Ethereum private blockchain	To manage access to IoMT devices and relevant medical data.	Access control, Traceability and log integrity, Data privacy.
[94]	1-IoT devices is subject to data theft, DDoS attack, hacking, remote hijacking 2-Data integrity and privacy issue	**On-chain:** • data collected from the body sensors after being filtered, • smart contracts that handle patient/doctor registration, provide access control and monitor patients.	Ethereum blockchain	To manage patients' information and medical devices.	Data integrity, Access control, Data privacy.

[95]	1-Centralized data storage	**Off-chain:** cloud server stores encrypted health data. **On-chain:** hash pointers to the storage location, metadata of the original dataset, and the transactional data.	Ethereum blockchain	To share personal continuous-dynamic health data in a secure and transparent manner.	Data integrity, Data sharing.
[96]	1-Assure the privacy, security, anonymity of dyslexic patients' data.	**Off-chain:** IPFS stores the multimedia payload. **On-chain:** hash of the media file and the final test results	Permissioned blockchain	To share Dyslexia diagnosis data securely with mobile medical practitioners around the globe.	Data integrity, Data sharing.
[97]	1-Security challenges brought by cloud computing.	**Off-chain:** cloud server stores the examination results. **On-chain:** indexes of the related healthcare data	Consortium blockchain	To reduce medical accidents in the telemedicine system by providing health data integrity & traceability.	Data integrity, Data traceability.
[98]	1-Security and privacy issues within patient monition healthcare applications/-platforms.	**Off-chain:** IPFS stores encrypted health data. **On-chain:** access control policies, pointer to data, registered patients, and added/removed devices.	Ethereum private blockchain	To help diabetic patients share securely their collected medical data with their physicians and help them control their data access.	Data integrity, Access control, Device & patient authentication.
[99]	1-Security issues within Mobile edge computing server, used to store the collected health data.	**Off-chain:** MEC server store collected health information. **On-chain:** collected health data after validation	Ethereum blockchain	To store collected health data securely.	Data integrity
[100]	1-Privacy issues when sharing data in real-time monitoring systems of diabetic patients.	**Off-chain:** cloud storage used to store data collected from wearable devices. **On-chain:** access permissions, log of the data accessed	Ethereum private blockchain	To share securely the collected personal information while allowing the patients to control their data access.	Access control, Log immutability & integrity
[101]	1-Remote patient monitoring healthcare data management issues at the level of access control, secure data storage, and user authentication.	**Off-chain:** the cloud and the data management module store the uneventful generated health data. **On-chain:** access grant transactions, legitimate healthcare provider and encryption of the eventful physiological data with its hash	Implemented their own blockchain	To share and store data securely while ensuring patient privacy. To record access log.	Data integrity, Data availability, Data sharing, Patient privacy.
[102]	1-Security and reliability issues in teleconsultation	**Off-chain:** IPFS stores the encrypted and anonymized medical files. **On-chain:** access control transactions, addresses of the stored medical data	Ethereum blockchain	To enable secure data sharing, and content-based access control to medical files.	Access control, Data sharing, Data integrity, User authentication.
[103]	1-No secure way to share collected data between health institutions 2-Centralized data storage 3-Give patients control over their health data	**Off-chain:** cloud database stores user health collected data, data request from healthcare providers & insurance companies, data access record/policy. **On-chain:** hash of collected health data, the access policies and access activity and request.	Hyperledger permissioned blockchain	To share securely collected health data between individuals, healthcare providers, and insurance companies.	Access control, Data integrity, Patient privacy, Log immutability & traceability.

Most challenges addressed in these selected papers are mainly related to the management of health data. Whether in a smart healthcare ecosystem or health information system, the vulnerabilities within the centralized design in current healthcare services, the security, privacy, and data access control issues prevent secure information sharing among peers. Hence, patients are unable to manage their fragmented data, which delays their diagnosis and treatment process. To address all these challenges, the blockchain has been introduced as a tool to:

- Share health data or diagnosis results in an efficient and secure way [88], [92];
- Manage access rights to the healthcare record or IoMT devices [75], [93], [104];
- Provide data audit and data provenance of shared records [67];
- Track patients' medical history [85];
- Manage cryptographic keys;
- Store collected health data such as: patients' medical questionnaire results [87] and validated health data collected from wearable devices [99].

A blockchain is a distributed ledger that records and shares all transactions that occur within the blockchain network [105]. It was introduced into the healthcare ecosystem since it can maintain data integrity, availability, transparency, reliability, immutability, resiliency, and traceability. Once the block is appended to the blockchain, the transactions can no longer be modified or deleted. We noticed that the role of the blockchain and its impact on the health system are mainly related to the information stored on the blockchain in transaction format and to the smart contracts' role in managing data. Smart contracts are self-executing scripts, deployed within a blockchain network, invoked when performing a transaction to execute the term of a contract/procedure on every node in the blockchain [105]. It can be utilized to:

- Monitor and report all actions performed on data thereby ensuring data provenance and auditing;
- Revoke access to the violated data, thereby preserving sensitive medical data against external attacks;
- Register all actions carried on data thereby ensuring accountability and creating a log for traceability;
- Verify the integrity and accuracy of the requested information;
- Handle registration of users (e.g., staff member of a hospital) or users' devices;
- Provide identity/device authentication and access verification;
- Define the access right of each of the system entities thereby ensuring the data access management in a flexible and secure manner. It could be designed to provide time-based or role-based access to data assets.

Additionally, research has shown that the blockchain cannot store large amounts of data due to its scalability issue. The scalability problem is related to the block size [106]. According to Dinh TT. *et al.*, the increase

in block size leads to a proportional decrease in block generation rate, thus hindering the overall throughput [43]. For example, in a bitcoin network, the block size should not exceed 4MB to get a maximum throughput of at most 27 transactions /sec [107]. Hence, storing the healthcare data on blockchain will impact its performance since a blockchain of hundreds of petabytes requires more computing power and more network bandwidth [90]. In fact, a majority of 90% of the studied papers addressed the blockchain limited storage capacity by adopting a combination of on-chain and off-chain storage. They stored the healthcare datasets (encrypted or not) off-chain, in a cloud repository, Interplanetary File System (IPFS), EHR databases, PostgreSQL, or institutional databases. As a result of having original data kept in insecure databases, what should we store on the blockchain, to share, handle, and access data in a secure way? To exchange health data while simultaneously maintaining the integrity and non-repudiation of the datasets, some authors proposed to store the hash pointers or the storage location URL on-chain. Others used the blockchain as a record keeper of keyword ciphertext to help users retrieve the requested data while protecting data security with searchability [68], [69]. Some stored the hash of the off-chain record on the blockchain to allow the verification of data authenticity in case a malicious database administrator altered data. Jiang S. *et al.* added the patient as well as the hospital signatures alongside the hash of the medical diagnostic report, so both parties cannot repudiate the administered treatment [69]. Likewise, Wang Y. *et al.* included in the data transactions the data provider's signature, in addition to the keyword ciphertext and data owner's account, in order to provide proof of the transactions' validity [68]. A digital signature provides authentication, integrity, and non-repudiation of a message and its source. Storing the data owner/provider signature in a blockchain transaction provides better signature preservation than certification authorities [108]. Once stored on the blockchain, the signature becomes a single shared source of truth, where all network actors/nodes can see the same signature, yet none can alter it [108]. To address the risk of data tampering when storing data in a cloud server, Wang S. *et al.* proposed to keep the cloud server signature on-chain thus entailing patients to verify the correctness of their data sent by the cloud [81]. Moreover, some researchers proposed to store the cryptographic key on-chain [65], [72], [81], [92]. Various encryption mechanisms are often required to handle privacy leakage and access control when sharing the health record stored in a cloud or any storage system. In such cases, relying on a third-party authority to set and distribute keys has disadvantages. The encryption keys, once discovered by any attack mechanism, expose the system to data breaches. Using the blockchain to maintain keys solves the single point failure problem of centralized key management, and it ensures that the management and distribution of keys is more secure [81].

While most papers adopted the on-off-chain model, some preferred to store the entire system's data on the blockchain while being watchful of any excessive storage. For instance, Alexaki S. *et al.* used universal interpretable code to represent the diagnosis, executed activities, and prescribed activities before storing them on-chain to minimize the storage requirements [74]. Previously conducted research has highlighted the fragmented health data along with centralized data stores, both of which hinder data sharing and jeopardize patients' privacy since any healthcare actor would be capable of illegally exploiting patients' data [61], [66], [68], [70], [75], [103]. According to HIPAA, patients have rights over their health records. We cite the patients' right to [109]:

- View or obtain a copy of their health data;
- Request errors amendment in their medical records;
- Control who can access their data and file a complaint in case of violation;
- Be informed about who accessed or updated their health data.

In fact, most frameworks placed the patients as the sole custodian of their medical records by having them handle the access control agreements. To guarantee their privacy, blockchain should not disclose the identity of the patients or data [65]. This could be achieved by:

- Storing de-identified health data on the blockchain [65], [68], [101]. For example, Pham H. *et al.* used anonymous accounts (EOA) to protect the identity of the patients [94]. In case any attackers gained data access, they will not be able to identify the patients' identity in real-life.
- Using an encryption scheme on the patients' data before storing them on the blockchain [65], [68], [83], [84], [92]. The storage of hash value or encrypted data prevents the risk of privacy leakage, as only authorized personnel can verify the accuracy of data or decrypt information when necessary.
- Providing a granular access control across health records. Data owners selectively share part of the record with the data requester based on different criteria [93]. Studies showed that the type of blockchain contributes to privacy preservation[62], [76]. In a permissioned blockchain, the network management and validation process are handled by a predefined group of participants, making it distinct from a permissionless blockchain, where anyone can join the network, read the blockchain's data, and participate in the consensus mechanism [78], [80], [103].

Chapter 3

Healthcare Supply Chain: Tracing the Drug Journey

3.1 Introduction

A supply chain ecosystem describes the processes that involve designing, engineering, manufacturing, and distributing products or services from suppliers to end-consumers [9]. Because these processes affect the goods, information and financial flows, some regulations are set to protect the consumers' right [110]. The eight basic consumer's rights recognized by the United Nations involve the right to safety, the right to be informed, the right to redress and the right to a healthy environment [10].

In the United States, the Food and Drug Administration (FDA) maintains consumers' right by promoting and protecting public health, through goods control and supervision. It works on applying predefined regulation and by protecting and promoting the development of human and veterinary drugs, biological products, medical devices and radiation-emitting products, human and animal food, and cosmetics [11]. Lack of real-time transparency and limited data provenance are two key challenges facing the supply chain, notably the drug supply chain [111]. They create dilemmas between the different supply chain stakeholders since there's no mechanism to track and authenticate drugs. Drug traceability is the ability to monitor and

audit pharmaceutical products as they move through the different supply chain stages. It ensures product safety and eliminates fraud or counterfeit drugs. Many incidents have occurred in the past years, ultimately putting to question the supply chain reliability and its product data accuracy. In January 2008, Baxter Healthcare Corporation recalled various lots of heparin, an anticoagulant medication, after associating the product with adverse events, including deaths [112]. More than three months later, U.S. FDA was able to establish a link between a contaminant found in heparin, a highly sulfated chondroitin sulfate, and the serious adverse events seen in patients given heparin. They traced back the contaminant to 12 different Chinese companies, and they found heparin batches shipped to 11 countries [113]. Table 3.1 provides information gathered from latest Food and Drug product recalls published on the FDA websites [12].

Table 3.1: Food and Drugs supply chain breaches reported by the U.S. FDA [12]

	Date	Brand	Product Description	Recall Reason
Food supply chain breach	03/08/19	Fullei Fresh	Organic Bean Sprouts	Listeria monocytogenes
	12/30/22	Full Circle Market Naturally Better	Oat Honey Organic Granola	Undeclared almonds
	01/13/23	Utopia Foods Inc.	Enoki Mushrooms	Listeria monocytogenes
	02/06/23	JSJ	Cake	Undeclared egg
Drugs supply chain breach	03/19/20	CVS, Rhinall, Humist, more	Nasal Products and Baby Oral Gels	Microbial Contamination
	06/11/20	Lupin	Metformin Hydrochloride Extended-release Tablets	Detection of N-Nitrosodimethylamine (NDMA)
	12/27/22	Hospira, Inc.	Vancomycin Injection	Presence of Visible Glass Particulates
	02/02/23	EzriCare Delsam Pharma	Artificial Tears Lubricant Eye Drops	Potential microbial contamination

Based on the collected information, we highlight the multiple breaches encountered daily on a supply chain [13], [14]. Not only is people's health affected, but businesses also undergo damage. The company with a recalled product will suffer a reputation loss due to negative publicity and will see it sales reduced dramatically [13]. During the investigation period, all related products will be affected, and some businesses shut down until the origin of this supply chain breach is detected [13]. Between the market expansion, the growth in suppliers' relationships, and the rising consumer demand, the supply chain complexity has increased and revealed the need to meet new challenges. The key objectives of the supply chain, including cost, quality, speed, dependability, risk reduction, sustainability, and flexibility, are not fully achieved [13]. Transparency and traceability need to be enhanced in manufacturing supply chains [14], [17], [57], [59], [114]. The main supply chain risk lies in the product journey. We need more knowledge about the product, its origin, processing, and shipping journey [14]. Consumers are unable to verify the integrity of the acquired product; they have to trust the certification logo printed on products. Verifying this certification integrity requires strenuous auditing. Transparency must be enabled not only to regain the consumer's trust but to help the producer get a better perspective of the supply chain breaches and understand how management product decisions and environmental circumstances can affect a product. Achieving transparency requires accurate data collection and secure data storage, a difficult task currently entrusted to third parties through centralized information depositories [14]. As previously highlighted, current pharmaceutical supply chains are centralized. Centralized supply chain management systems expose the supply chain to corruption, fraud, and tampering.

Blockchain has emerged as a new distributed information technology; it represents a new approach in supply chain area, where visibility and transparency of product flows are the principal challenges. However, introducing blockchain into the supply chain ecosystem introduces new challenges, particularly at the storage level [59]. According to Zheng Z. et al, blockchain is not always sufficient for storing data [44]. With the increasing number of transactions, the blockchain has become heavy. Hence scalability becomes challenging. For example, in a Bitcoin network, the block size is limited to 1 MB, and a block is added

every 10 minutes [44]. Transaction's rate is limited to seven transactions per second, which is not enough for the trading system. Increasing the block size will reduce the network efficiency.
To overcome the blockchain scalability issue, Feng Tian, propose to integrate BigchainDB into the supply chain ecosystem. As proposed by McConaghy *et al.*, BigchainDB combines the key benefits of distributed Databases - high throughput, low latency and high capacity- with the key benefits of blockchain - decentralization, immutability, creation and movement of digital assets [57].
Based on the above studies some questions have emerged:

- What are the benefits of introducing the blockchain to the supply chain?
- Can we trust the information shared in a supply chain traceability system?
- What are the challenges we need to address when integrating the blockchain in a supply chain?

3.2 Methodology and case study

To evaluate blockchain-based supply chain efficiency and sufficiency to create a reliable, transparent, authentic, and secure system, we have adopted the theory built based on case studies as a research strategy. Working on real cases will highlight the challenges and characteristics to be taken into consideration in order to build an efficient blockchain-based supply chain. To confirm the theoretical study, knowledge of the practical and real-world application of the blockchain in a supply chain ecosystem is needed. By theoretical study, we imply study not deployed on a large scale. According to Eisenhardt k., case studies emphasize the rich, real-world context in which the phenomena occur [18].
Several startups have already identified the blockchain as a new paradigm that aims to enhance supply chain management. We summarize, in Table 3.2, the main goal of the most prominent supply chain implementations and compare them according to the blockchain type and tracking system used. By introducing blockchain into their supply chain, these startups aimed to track, record and verify goods as well as protect them from fraud and tampering.

Table 3.2: Blockchain-base Supply chain start-ups

	Main Goal	Blockchain type	Tracking system
Ambrosus [19]	Ensuring the origin, quality, compliance and proper handling of food and pharmaceutical tracked product.	Public: Ethereum Blockchain Private (for testing): Ambrosus Blockchain	Tag : QR code Tracer Sensor : Biosensor
Ascribe [115]	Web-based solution, to track, record and verify ownership, in the digital art market. All the digital contents are securely shared with artist and clients.	Public: Bitcoin Blockchain	SPOOL protocol: Used for time-stamping evidence of ownership transactions.
Blockverify [116]	Identify counterfeit goods, stolen merchandise and fraudulent transactions by introducing blockchain into the supply chain. Used for luxury and pharmaceutical items.	Public: Bitcoin Blockchain	Block Verify tag
Chronicled [117]	Protect goods from fraud and tampering.	Public: Ethereum blockchain Future work: implement their own private blockchain	IoT devices: such as temperature logger, tamper proof smart tag.

Table 3.2 – *Continued on next page*

Table 3.2 – *Continued from previous page*

OwlChain [118]	Build a trusted ecosystem between the producer and the customer, by using public and transparent information. Mainly adopted in the food industry.	Private: AMIS blockchain based on the Ethereum technology.	Tag
Provenance [119]	Tracing back and verifying the origins, attributes and ownership of a specific product.	Public: Ethereum Blockchain	Tags: QR code, NFC tags, Laser-engraved barcodes, 3D scanning
Modum [20]	Track and trace pharmaceutical products in a secure way that meets all the requirements imposed by GDP [1].	Public: Ethereum Blockchain	IoT sensor devices, and QR code
Everledger [120]	Tracking and protecting valuable assets (such as: diamond) from fraud, trafficking and theft.	Public: Ethereum Blockchain Private: Hyperledger Blockchain	Thumbprint
Verisart [121]	Certifying, documenting, verifying and tracking artwork ownership.	Public: Bitcoin Blockchain	Image identification algorithm

Table 3.2 – *Continued on next page*

[1]Good Distribution Practice regulation used in the European Union

Table 3.2 – *Continued from previous page*

TrustChain [122]	Tracking and authenticating Jewelry such as diamonds.	Public: IBM Blockchain based on the Hyperledger Fabric	recording on the blockchain ledger: high-resolution photos of each diamond at every touchpoint along its journey, certificate of authenticity and product details

As shown in the previous section, integrating the blockchain into the supply chain ecosystem brought significant new challenges notably on the blockchain level. To build a blockchain-based supply chain management, we need to take into consideration not only the blockchain technology but also the reliability of collected data.

To study how startups, integrate the blockchain into their supply chain ecosystem, we need to understand first their tracking system and the blockchain's role in their platforms' architecture. The efficiency and sufficiency of their blockchain-based supply chain will be developed in the discussion part. We selected Ambrosus and Modum as real cases to study since we could obtain sufficient information and they are related to the food and pharmaceutical supply chain: our main interest. By combining the theoretical findings with those drawn from real cases, we are able to address the emergent questions listed above.

3.3 Description of the selected cases

Ambrosus and Modum two Swiss Startups, have developed a system that merges IoT, blockchain technology and real-time sensors to trace and transmit products' information during the whole manufacturing

process. They aim to optimize supply chain visibility and quality assurance. Modum specializes in the pharmaceutical supply chain to ensure the safe delivery of pharmaceutical drugs in compliance with the GDP requirements. Ambrosus specializes in food and pharmaceutical supply chain to ensure the quality and safety of product consumption. For each case, we will describe the tracking system and the blockchain integration into these startups' system.

Case 1: Ambrosus

1. Tracking system

   The Ambrosus network uses tags, tracers and sensors to track products throughout their life cycles. Their goal is to associate the product with the packaging and the transportation car, in a way that, if compromised, a notification is sent to the blockchain [123]. Tracking components are customized according to the product type and based on the clients' needs. To track a fish from hook to fork, different tracking components are deployed [123]. A smart gel tag is applied at the surface of the fish to assure product authenticity because the gel will react to fraudulent manipulation. A container sealed with a sensor contains all the collected fish. The sensor will assure the integrity of the product since it will detect any opened container. Another sensor is added to check the temperature and the GPS movement during the shipment. A Charge-coupled device (CCD) camera can be introduced to record all occurred activity until product shipment. All these sensors are bonded together and then bonded to the QR code. A QR code is a matrix barcode that contains information related to the product to which it is attached. All data obtained from the product, related to the QR code rectification and collected by the sensors aggregate to the QR code [19].

   In the Ambrosus network, all tracking devices are authenticated by a public-private key cryptography. The sensors and the QR code sign the collected data before sending it to the edge gateway using RFID technology [19]. The gateway is a device composed of a microcontroller. It selects the received data before sending it to the blockchain through General Packet Radio Service (GPRS) technology [19]. Note that the edge gateway needs to operate for months; thus, it must be powered by batteries or power harvesting [123]. Once the received data is verified, it is saved on the blockchain. An Amber token is introduced to the ecosystem, and each amber will remain bonded to the product until a defined expiration date, such as a purchase [123]. A customer can download all the required data concerning his purchased product from the web application which is built on top of the application programming interface (API). The API is linked to Ambrosus's data storage infrastructure; hence

data becomes reachable to anyone who needs to verify the authenticity of his product.

2. Blockchain integration

   In the Ambrosus Network, Ethereum blockchain is introduced to verify products' quality based on predefined requirements and to verify the tracking devices' identity. Supply chain automated governance and data management is mainly related to the deployment of two smart contracts: the requirement smart contract and the measurement smart contract.

   The requirement smart contract defines the quality standards a specific item needs to maintain during the whole shipment till the delivery. The measurement smart contract stores:

   - the collected attributes for a given batch at a specific point along the supply chain [19]
   - the defined list of Ambrosus-certified measurement devices [19]
   - the root hash of the Merkle tree

   The list of statements defined in the requirement smart contract will be compared to the content of the measurement smart contract. Meeting all the requirements ensure that the shipped product remains safe and in good quality [19]. Measurement and requirement smart contracts are stored and are publicly available on the Ethereum blockchain.

   However, Ethereum blockchain has a limited capacity in handling large quantity of data. Putting sensors' collected data on the blockchain will degrade its performance because the blockchain can handle a limited number of transactions per second [123]. Thus, Ambrosus introduces InterPlanetary File System (IPFS), a distributed storage, alongside the blockchain, to store all sensors' data [19].

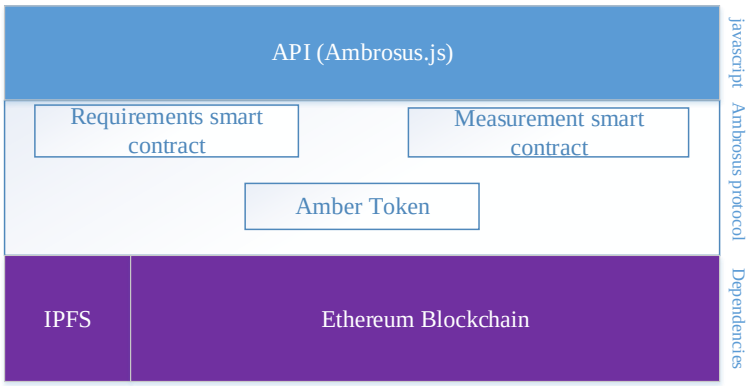


Figure 3.1: Ambrosus system

In the Ambrosus layered architecture, the Ethereum blockchain and the distributed storage are located at the lowest level (first layer) as shown in Figure 3.1. They represent the core of the Ambrosus system [123]. Parity is the programming language used to build smart contracts running on Ethereum blockchain, and because running transactions on the Ethereum blockchain became expensive, an Ambrosus blockchain was adopted as the main transactional network. The Ambrosus blockchain is written in solidity and built upon the Ethereum blockchain.

The second layer consists of the Ambrosus protocol. The main three components of the Ambrosus protocol are: the measurement repository, the requirement smart contract and the amber token [19]. All smart contracts associated with the Ambrosus protocol will run on the Ambrosus blockchain which will be copied to the Ethereum main network for further validation [19].

Above the Ambrosus protocol layer, we have the API, also known as the JavaScript layer. It allows developer to create and run Ambrosus contracts and objects in the Ambrosus platform without any blockchain programming knowledge [19]. Developer can use JavaScript or html to connect their own hardware to the Ambrosus network.

Case 2: Modum

1. Tracking system

   In the Modum network, monitoring begins in the web/mobile app, where setup, review and reporting happen. A quality manager creates a shipment profile with monitoring criteria and program notification to alert the team to any problem. The deviation can be visualized on the dashboard. The logistic teams activate the logger, also known as SensorTag, using an NFC plate, and connect it with the shipment ID. A smart contract is created to each shipment.

   The SensorTag is used to measure the environmental conditions that the shipment is subject to, store collected data in its internal memory and send data to the mobile application [124].

   Barcodes are used to identify the items handled in the Modum ecosystem. A sensor tag is associated with a unique MAC-Address through a QR code, and a packet is associated with a unique track-and-trace number through a distinct QR code. The camera of the Android device associates a logger with a shipment by capturing the QR code of both the sensor and the packet. The track-and-trace number/MAC-address association is sent to the server or is saved on the sensor internal memory if

the server is not available [124]. The logger starts recording the temperature using the setting from the shipment profile. The temperature is recorded every 10 minutes in the sensor's internal memory. The server stores the track-and-trace number/MAC-address association, broadcasts the smart contract and stores the smart contract ID on the sensor device. When the client receives the packet, he will scan the track-and-trace number and request the temperature measurements downloaded from the sensor via the Bluetooth Low Energy (BLE). The data is sent to the smart contract to verify the compliance and a report is sent back to the client's mobile application [124]. There is no need to open the package to perform the checks [20].

2. Blockchain integration

   In the Modum Network, Ethereum blockchain is introduced to verify products' temperature compliance with GDP regulations. After the data verification process, smart contract stores, on the blockchain, the measurement hash and the uniform resource locator (URL) that points to the actual measurement data stored in the PostgreSQL [124]. Raw temperature data and user credentials are stored in PostgreSQL because collected data is too large or too sensitive to be stored on the blockchain.

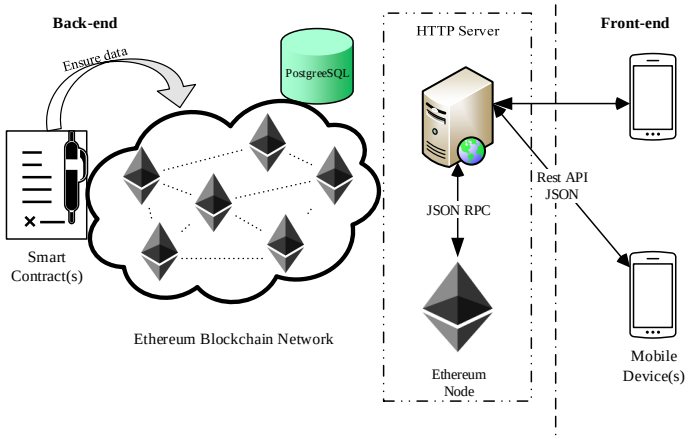


Figure 3.2: Modum system [20]

The blockchain is in the back end of the Modum system, next to the server and database as shown in Figure 3.2 . The role of the server is to create, call and modify smart contract. This is done through an Ethereum node hosted by the server. HyperText Transfer Protocol (HTTP) server communicates with the Ethereum node over JavaScript Object Notation (JSON). Written in solidity, smart contracts run in an Ethereum virtual machine (EVM) to ensure that the shipment complies with the temperature

required by the GDP. For each medical product type, hence for each shipment, a smart contract is configured [124]. A shipment-specific smart contract contains the temperature logger ID, the shipment ID and the alarm criteria. If the temperature data collected by the sensor does not meet the GDP regulation, the sender and receiver are notified to deal with the issue [20].

The communication between the front end and the blockchain is done through HTTP server over Representational State Transfer (REST) API, using JSON to encode and decode requests and responses [124].

3.4 Discussion

Different tracking components are deployed in Ambrosus and Modum systems, ranging from sensors to tags and tracers. The selection of the appropriate tracking device relies mostly on the product. For example, some products such as meat, fish, or vaccine must be maintained at a specific temperature and humidity conditions, requiring environmental monitoring sensors [123], [124]. Alcoholic beverage industries must ensure product authenticity through products' lifecycle, and this process requires the usage of unique tags. Applied to corks these unique tags change color if the cork is removed or a needle is inserted to extract or tamper the liquid.

Chemicals and biological sensors are implemented to enable freshness investigation of food products and to assess food or medicine adulteration, authenticity and toxicity [125]. Biosensor, an analytical sensor, is introduced to inform the interested supply chain parties about the biological content of a given product. Detecting the presence of allergens in food products such as milk, soybeans, eggs, peanuts, etc., is now a real concern since the prevalence of food allergies due to trace amounts of allergens is increasing [125]. Biosensor relies on the biomolecule's recognition properties such as enzymes antibodies to monitor the product, through a variety of methods including colorimetric and mass-based detection [125].

The pilot project led by Modum showed the importance of having offline features, at the level of the tracking devices level, where data is stored internally until it can be uploaded on the blockchain [124]. This could provide the tracking system a more robust quality.

As observed in Table 3.2, multiple tracking devices are sometimes recommended to meet all the product tracking requirements. In their paper, Mackey *et al.* also highlight the same point and indicate how the

most mature digital anti-counterfeit technologies include mobile and RFID-based technologies in order to enable fake drug detection, authentication and tracking [126].

Data collected from the tracking devices are processed and sent to the storage system or blockchain. Ambrosus and Modum adopted different technologies to transfer data. We can mention RFID, BLE, NFC, GPRS or 3G. Table 3.3 shows different criteria used to compare the communication protocols. According to Al-Sarawi *et al.*, each protocol has benefits and limitations; the most suitable communication protocol is selected according to the application needs [127].

RFID, BLE and NFC offer low power consumption and low setup time, but the maximum data rate for NFC is 424 Kbits per second, which is unsuitable for transferring large amount of data. It's not the case with RFID, which has the highest data rate of 4Mbps. In comparison, a widespread mobile network like 3G provides reliable high-speed internet connectivity, efficient for continuous streaming; however, 3G has a high-power consumption profile, making it unsuitable for local network communication [128].

On the security level, we can identify multiple vulnerabilities for each protocol, that we must take into consideration during the implementation. For example, a device, if not encrypted, can be vulnerable to BLE attacks, such as replay attack or fuzzing attack [129]. Collected and transferred data must be safe and reliable before being stored in an immutable decentralized database. Protecting data from hacking, fraud and manipulation during the whole supply chain process is crucial because the product's quality depends on it.

Ambrosus and Modum platforms have integrated security measures into their traceability system to guarantee integrity and authenticity of the handled data.

In Ambrosus, all authenticated devices sign data before transmitting it. The devices' signatures are verified before recording any message onto the blockchain. To verify the identity of a device, we can check the list of authorized devices publicly available in a smart contract. A measurement is ignored if it is sent from a non-authorized device. Besides, if a device becomes compromised or faulty, it will be disconnected from the system [123].

In Modum, all components with a serial number are registered in the database as authorized devices in order to identify forged, tampered or stolen loggers [20]. The sensor housing is tamper-resistant and waterproof; thus, it could not be physically disassembled or manipulated. All data, such as the measurements and timestamps, are signed by the logger before being transmitted. This will guarantee an end-to-end authenticity [20]. The private key is shielded by a cryptographic co-processor. Plus, Modum system offers a restricted access control, where only the authorized users can interact with the loggers.

Table 3.3: Comparison between Communication Protocols

	RFID	NFC	BLE	Cellular	
				GPRS	3G
Set-up Time	< 0.1 msec [130]	< 0.1 msec [130]	Low	N.A.	N.A.
Power consumption	Ultra-low power [127]	50 mA Low power [127]	30 mA Low Power [127]	High power Consumption [127]	High power Consumption [127]
Data Rate	4 Mbps [127]	106, 212 or 424 kbps [127], [130]	1Mbps [127]	9.6- 172 kbits/sec [131]	7.2 or 56 Mbps
Communication Range	Short Range Up to 200 m [127]	Short Range 0-1m [127]	Short Range 100 m [128]	Several km [127]	> 5km [128]
Cost	Affordable [130]	Low [130]	Medium [130]	Expensive [130]	Expensive [130]
Security vulnerabilities	tags face the risks of being broken, cloned, counterfeited and distorted [132] security issues on wireless communication links security risks within readers [132]	Subject to security attack during transmission (ex: man, in the middle attack, eavesdropping, data corruption, data modification) [130], [133]	Susceptible to DoS attacks Fuzzing attacks and eavesdropping attacks [129]	Vulnerability in authentication procedure absence of a mechanism that ensures data integrity [134]	Wireless link threats Network services threats Terminal threats [3g]

Principal Application domains	Industry (Tracking, Inventory, Access) [128]	Smart Cities, Industry, Secure transactions (payment) [133]	Healthcare, Wireless headsets, Audio Applications [128]	Local Network (M2M) [128]	Smart Cities, Smart Building, Automotive [128]

Once building an effective traceability system, we need to study the efficiency of the blockchain technology to store and manage all transactions that occurred in the supply chain ecosystem. As shown in the Ambrosus and Modum platform, blockchain was integrated into their ecosystem to improve the tracking system and data management. Blockchain reduces fraud, errors, and delays identified in the actual supply chain ecosystem. It increases the trust between the customer and the supplier through the distributed ledger that is updated and validated in real time with each network transaction. In the tracking system, blockchain was integrated to:

- Provide transparency, reliability, and integrity of the products' data collected throughout the entire lifecycle.
- Provide tracking device authenticity

In the data management, blockchain was integrated, to ensure availability, accuracy, and accessibility of data for all supply chain actors. Blockchain will improve business decisions and give deep insights into all the system's vulnerabilities.

Despite all these advantages, blockchain showed some limitations. In fact, Ethereum blockchain, adopted by both platforms, has a limited capacity in handling a large quantity of data. Ambrosus and Modum introduced a distributed file system: IPFS [19] and an object-relational database system PostgreSQL [20] respectively, besides the blockchain, to store large data.

Using IPFS depends on the product's type and the client's demand. For products requiring a high level of security such as pharmaceutical products, IPFS is no longer used, because it's not secure. Plus, it fails in providing cumulative analysis and flexibility in handling interconnected data. IPFS will be replaced by the Ambrosus's own storing system: Ambrosus Blockchain. Ambrosus blockchain is implemented to enhance

system performance by avoiding throughput degradation and latency faced in Ethereum blockchain.
According to D. Tien Tuan Anh *et al.*, blockchain systems are not ready for mass usage [43]. Following their comparative study of Ethereum, Parity and Hyperledger, where they used their BLOCKBENCH framework, they reached different conclusions. Among those findings, we highlight the most interesting ones in our case:

- In terms of throughput, Hyperledger performs the best. Compared to Ethereum, the gap is related to the adopted consensus protocol adopted. Hyperledger uses Practical Byzantine Fault Tolerance (PBFT), where the communication cost of broadcasting messages is cheaper than Proof of Work consensus protocol adopted by Ethereum.
- In terms of scalability, Parity performs best, due to its constant transaction processing rate. This is not the case with Ethereum and Hyperldeger, whose performance is affected by the number of used servers. In fact, Hyperledger will stop working when the number of servers and nodes reach a certain threshold because the number of dropped consensus messages will increase due to channel request congestion. In Ethereum, the consensus protocol is computationally based. Hence, increasing the number of servers and nodes will lead to throughput degradation since the computation difficulty has increased with the increase of the network's size in order to avoid long propagation delays.
- In terms of crash failures, Ethereum and Parity are both unaffected. This does not apply to Hyperledger, where the consensus protocol PBFT cannot tolerate more than 4 failures in a 12-server network.
- In terms of security attack, Ethereum and Parity are both vulnerable. The vulnerability is related to the consensus protocol adopted. For example, in Ethereum, the Proof of Work consensus is probabilistic. Hence, two blocks can append at the same time, creating a fork, exposing the system to double spending attack. It is not the case with Hyperledger where the Practical Byzantine Fault Tolerance consensus is considered safe with no forking problem.

Caro *et al.* implemented a fully decentralized traceability system for the Agrifood supply chain relying either on the Ethereum or the Hyperledger Sawtooth blockchain implementations. Based on their practical test, the implementation based on Hyperledger Sawtooth showed better results compared to the Ethereum one, in terms of latency, network traffic and CPU load. However, Ethereum is more advantageous in term

of scalability, reliability and system maturity, enabling a large number of participants [114].

As revealed earlier, to deal with the blockchain limitations, some proposed to introduce a storage system next to the blockchain; some implemented their own transactional blockchain while others adopted a blockchain technology with some database functionality like BigchainDB [135]. Various blockchains are now publicly available and ready to be implemented such as Ethereum [136], Hyperledger Fabric [137], [138], Hyperledger Sawtooth [137], [138] and BigchainDB [135]. It's important to consider the blockchain's different properties (decentralized control, immutability, creation, and movement of digital assets) and capabilities (throughput, latency, capacity, scalability) before choosing a blockchain implementation over the other.

According to Baliga A., the security of a blockchain based system is related to the security and the robustness of the adopted consensus model [139]. The consensus protocol forms the core and the working entity of blockchain [46]. A bad consensus mechanism can compromise the data recorded on the blockchain. If the consensus mechanism fails it will lead to issues such as blockchain fork, consensus failure, dominance and cheating [139]. An efficient consensus protocol implementation can enhance economy growth, by ensuring the proper functioning of the blockchain and by avoiding any blockchain architecture malfunction [46]. The security of a blockchain could be related to the blockchain type. Using Ethereum blockchain means that ledger can be viewed by anyone connected to the network. However, stored data can be sensitive; a certification authority is required to control the supply chain actors' role (read/write access) and identity. Data can be accessible to some supply chain members/stakeholders and limited to others. Hence, it's important to choose the blockchain type we want to adopt in our ecosystem, the one that will help us achieve the supply chain's main goal.

3.5 Conclusion

The blockchain is introduced to achieve the supply chain's objectives, by reducing the risk emerging from the tracking system and data management.
Deploying blockchain in the supply chain ecosystem brought many benefits, notably:

- Creating more transparent and accurate end-to-end tracking;
- Increasing trust between the producer and consumer, by improving visibility and product compliance with international standards;
- Reducing paperwork and administrative costs;
- Reducing or eliminating fraud and counterfeit products;
- Facilitating origin tracking;
- Recalling a product in a time-efficient way.

However, integrating the blockchain into the supply chain ecosystem brought important new challenges notably on the blockchain level. We need to consider the properties and capabilities of available blockchain implementations before choosing the most suitable blockchain to such an ecosystem. To build a blockchain-based supply chain management, we need to take into consideration not only the blockchain technology suitable to our business but also the reliability of collected data.
Storing reliable information requires a reliable interaction between the blockchain and all ecosystems' constituents (These consists of tracking devices and actors).
To build a blockchain-based supply chain, we need to consider these requirements:

- Select a blockchain according to different key criteria notably: Throughput, latency, capacity and scalability (A multi-criteria decision-making can be applied to choose the most suitable blockchain to our deployed ecosystem.)
- Implement a dual storage architecture to handle large amount of data, without degrading the blockchain performance (An additional private blockchain could be introduced to the system architecture.)
- Choose the tracking devices based on the main product criteria we want to track or monitor.
- Choose the communication protocol based on the speed, data rate, communication range, power consumption, cost or any criteria deemed essential in the supply chain environment.
- Try to fill the security vulnerabilities found in the communication protocol to provide a secure and reliable traceability system.

- Create a secure tracking environment beginning by authenticating the system tracking devices and making sure all transferred or collected data is encrypted and signed.

Chapter 4

Healthcare Delivery Chain: Tracing the Patient Journey

4.1 Introduction

In today's healthcare delivery system, the patient is at the center of a complex network in which every interaction between the different stakeholders contains opportunities for error. Information may be collected, recorded, or communicated inaccurately. In fact, patients may visit multiple medical institutions for consultations or treatments. As a result, their medical data become fragmented across these institutions. Data fragmentation hampers the quality-of-care patients receive and incurs high healthcare costs. In order to enhance the outcome of healthcare services, medical data require proper management. Therefore, to address the data management challenges in the healthcare sector, we adopted a case study approach to examine a real-life issue and propose a solution that enables patients to personally manage their health data.

In Lebanon, adverse drug events are mainly related to fragmented care caused by the absence of a well-established national health information system and the inefficient cooperation between the different health entities. Consequently, these aspects affect health delivery and delay patients' treatment [21], [22].

In addition, inadequate communication between patients and physicians, especially during emergencies, can increase the former's chance of drug interactions and allergic reactions [140]. A cross-sectional study performed by Ramia E. *et al.*, on the knowledge and practices of Lebanese outpatients regarding their medication use and risks showed that of the 921 studied patients, 38.7% of the patients do not share their ongoing medications each time they visit their physicians [140]. The study also reports that physicians do not regularly assess medication history or ask about previous adverse drug reactions before prescribing new drugs [140].

This work addresses the challenge of minimizing adverse drug reactions in the Lebanese healthcare system that are caused by the intake of inappropriate medication. This can be achieved through designing a system that stores patients' prescriptions and drug allergies and enables sharing this information with physicians in order to alert them of any unsuitable medicine based on the collected data. To achieve our main objective, we propose to build a blockchain-based healthcare system powered by Semantic Web technology.

In recent years, blockchain emerged as a distributed information technology to eliminate the need for third-party intermediation and to overcome the security issues related to a centralized ledger [28]. Several studies proposed the blockchain as a potential solution to alleviate some issues associated with healthcare scenarios. Blockchain provides a permanent record of transactions ordered into an immutable block. It maintains data integrity, immutability, availability, resiliency, and traceability [28].

At another level, the World Wide Web Consortium brought the Semantic Web as a tool to facilitate the management and sharing of knowledge between systems [141]. Semantic Web technologies may address interoperability and data interpretation challenges among healthcare information systems by promoting interoperability standards and proposing ontologies to represent knowledge within the healthcare domain [142]. In our context, Semantic Web technologies seem like a potential tool to enable knowledge inference and discovery.

To achieve our goal, we first provide a comprehensive review of the existing work related to managing patients' prescriptions to protect them from adverse drug events. Furthermore, we will review the use of Semantic Web technologies in the healthcare delivery system since the Semantic Web is part of our solution.

Next, we describe our proposed solution that combines Semantic Web and blockchain technologies. Finally, we conclude with our challenges and future work.

4.2 Overview and related work

As per the U.S. Department of Health and Human Services, adverse drug events (ADEs) can happen anywhere in an inpatient or outpatient setting. ADEs account for 1 million visits to the emergency department and approximately 100,000 hospitalizations [143], [144]. In addition, hospitalizations due to ADEs are far more common among the elderly, accounting for one in every six hospital admissions [145]. Electronic prescriptions were introduced to enhance patient safety by reducing medication prescription errors mainly due to drug-drug interactions, drug allergy, missing or inappropriate dosing, illegible or unsigned orders [23], [24]. According to Ammenwerth, E. *et al.*, electronic prescribing reduces the risk for ADEs from 30% to 84% since it reduces the risk of medication errors [24]. Furthermore, including advanced clinical decision support (CDS) in electronic prescribing systems minimizes the incidence of adverse drug events even more [24]. Nakhla, Z. *et al.* proposed a system that helps doctors in the diagnostic process and detects adverse drug events in prescriptions using ADE ontology and IoT [146]. They defined an ADE ontology that includes all the concepts and instances of ADE along with the relationships between them. They used their ontology alongside the data collected from the patient (notably: symptoms, antecedents, and taken drugs) to construct a personalized ADE ontology for each patient [146]. Apart from the advantages of the e-prescription system, various concerns were highlighted, including security and communication. Patients' inability to acquire their drugs prescription before reaching their pharmacists is one of the communication hurdles identified by researchers. As for security concerns, e-prescription systems are susceptible to attacks, compromising patient privacy [25]. Aldughayfiq, Bader, et Srinivas Sampalli evaluated the architecture and digital security of the e-Prescription systems in eight countries. According to their research, most adopted systems are centralized, making them open to security concerns [26]. Adopting a decentralized system preserves patient privacy and reduces attacks on stored medical information. They also highlighted the importance of making the patient medication history available to all parties involved in the e-prescription system. This feature is not available to prescribers in some systems [26]. Most caregivers obtain medication history information while interviewing the patient during the

administration process, making information untrustworthy because it relies on patients' memory [26]. In addition, 91% of prescribers believe that generating alerts about the prescribed drugs based on multiple factors such as patient allergy history, current health condition, or previous cases of drug interactions will enhance the safety of medication prescribing [147]. Likewise, to improve the treatment process, Drug-Drug interaction alerts should be integrated into the e-prescription system [26]. Drug-allergy alerts and drug-drug interaction alerts are some features of the CDS. However, in some countries, the CDS is not part of the e-Prescription system. They are either unavailable or integrated into other healthcare systems, such as electronic health record (EHR) [26].

Due to the growing interest in using distributed ledger technologies in the healthcare sector, different research papers have introduced blockchain technology as a tool to alleviate some major challenges identified in medication management and protect patients from any adverse drug events. Thatcher, C. *et al.* introduced the blockchain to the e-Prescription system to control and monitor the prescribing of opioids as pain relievers to prevent the reoccurrence of any opioid crisis, avoiding overdose deaths among patients [148]. To enhance allergy information management, Ngassam, R.G. *et al.*, designed a blockchain-based mobile application to report and share allergy information between patients and healthcare professionals [149]. Mitchell, Ian, et Sukhvinder Hara proposed the blockchain as a tool to audit and report medication administration records to improve health services and safeguard vulnerable adults [150]. Li, P. *et al.*, introduced the blockchain to track and manage patients' prescriptions [151]. Garcia, R. D. *et al.* proposed an e-prescription system using smart contracts to reduce costs and scams [152]. Blockchain has been introduced into the healthcare ecosystem to maintain data integrity, availability, transparency, reliability, immutability, resiliency, and traceability. However, blockchain technology is not enough to tackle all the healthcare issues, especially in terms of interoperability and description of healthcare domain-specific concepts. Hence, we integrated the Semantic Web technologies into our healthcare ecosystem to provide semantic interoperability and enable knowledge inference and automated decision-making. We propose a conceptual model of a prescribing system, where we combine the Semantic Web and blockchain technologies to enhance drug prescription management and protect the patient from drug side effects or drug interactions.

Before outlining our solution's design, we will conduct a comprehensive review of the role of the Semantic Web in improving the healthcare delivery system since it is a crucial aspect of our solution along

with blockchain technology.

4.3 Semantic Web applications in healthcare delivery system

To explore the key research priorities at introducing the Semantic Web into the healthcare delivery chain, we searched the IEEE Xplore research database for the exact keywords: "Semantic Web" AND "Healthcare". We studied all the papers published between 2017 and 2020 and selected those which proposed a patient-centered framework integrating the Semantic Web as a tool to address healthcare issues. We excluded the papers related to the semantic deep learning approach, semantic modeling and analysis for the natural language process, and machine learning since they are not directly related to the proposed solution in this chapter. We reviewed the following questions:

- What healthcare issues do the authors address in the paper?
- How was the Semantic Web introduced into the healthcare system?
- What has the Semantic Web brought to the healthcare ecosystem?

The extracted data are reported in Table 4.1.

Table 4.1: Data extraction from the articles related to the Semantic Web applications in healthcare ecosystem

	Semantic Web applications in healthcare			
	Healthcare Challenges	Semantic Web		
		Component	Role	Impact on system
[153]	Lack of coordination between doctors due to inefficient healthcare data sharing system.	**SW language:** OWL	Used ontology to represent real-time information	Make the analysis process of real-time healthcare data more pertinent and relevant
[154]	Identifying the nearest hospital, ambulance department, pharmacy for a patient in a situation of emergency.	**SW language:** OWL **Query language:**SPARQL **Ontology editor:**Protégé	Used ontology to describe the rules, relations between hospital entities.	Enhance the efficiency of the emergency system by allocating services dynamically based on the incident severity and service availability.
[155]	Interoperability issues between different IoT healthcare systems.	**SW language:** OWL **Query language:**SPARQL **Ontology editor:**Protégé	Used the "pharmacy. owl" and "clinic.owl" created ontologies to represent knowledge about the pharmacy and clinical healthcare system respectively. Mapped the local ontologies into one global ontology to enable collaboration between healthcare systems.	Enhance interoperability between the clinic and pharmacy healthcare systems. Reduce delays and error in treatment by converting data into useful, understandable information, and actionable knowledge.
[156]	Interoperability issues in pervasive computing applications due to data heterogeneity collected from various sensors.	**SW language:** RDF **Semantic annotation model:** SSN ontology	Used SWT to annotate sensor data based on the SSN ontology.	Achieve sensor interoperability through a unified representation of sensor data that can be shared and reused among various pervasive applications.

[157]	Clinical Decision Support System (CDSS) challenges: • unstructured EHR data which may contain hidden risk factor • Processing heterogeneous data	**Ontology editor:** Protégé **SW language:** SWRL	Used semantic rules to translate the relationship between risk factors. Used UMLS medical ontologies to extract the risk factor concepts from the EHR.	Enhance the extraction of the risk factors from heterogeneous data, hence improved the prognosis/ diagnosis in CDSS.
[158]	Analyze and interpret extensive health data collected from heterogeneous medical connected objects.	**SW language:** SWRL	Used the "healthIoT" created ontology to represent the medical connected objects and their data. Defined SWRL rules to manage and interpret collected data.	Enhance semantic healthcare data interoperability between heterogeneous medical connected objects and alleviate the decision-making of doctors.
[159]	Query, integrate, and store efficiently the big dataset collected from heterogeneous health data resources.	**SW language:** SWRL **Query language:**SPARQL **Semantic reasoning framework :** Apache Jena **Semantic RDF repository:** Virtuoso	Defined MyHealthAvatar H-event ontology to map the NoSQL model to become knowledge. Defined SPARQL queries and SWRL rules for semantic reasoning to link patients' activity with physiological symptoms.	Enhance the knowledge management of large integrated datasets from heterogeneous data collections. Enhance the people's health knowledge discovery by providing an accurate health condition analysis.
[160]	Interoperability issues in IoT-based e-healthcare services due to the heterogeneity of collected data that makes challenging the extraction of understandable knowledge.	-	Used ontology to semantically annotate the collected data. Used semantic rules to detect the real-world situation to provide knowledge to the service layer.	Provide semantic interoperability between system components to help understand and infer the user situations from the real world hence enhancing depressive disorder assistance and care.
[161]	1. Interoperability issue of: • Medical devices between them • Medical devices with the health information system 2. Security and data privacy issues 3. Data integration issues, due to unstructured data	**Semantic annotation model:** Semantic medical devices (SMD) ontology & Unified medical language system ontology (UMLS)	Used SMD ontology to annotate data from sensors & medical devices. Used UMLS ontology to annotate medical documents & data from any medical information system. Used semantic reasoner to interpret the acquired medical data.	Facilitate the integration of massive and heterogeneous medical data. Enhance medical data interpretation. Improves decision-making for professionals.
[162]	Interoperability issue in IoT based healthcare system due to the heterogeneity of collected data.	**SW language:** OWL	Used Ontology to represent the healthcare data.	Enhance data interoperability where the ontological representation promotes better data handling and decision-making.
[163]	Low diagnostic accuracy in current symptom checker applications, because they: • rely on manually constructed knowledge models, • have a limited terminology process, • disregard user's health in the diagnosis.	**SW language:** OWL **Ontology editor:** Protégé **Semantic reasoner:**TrOWL	Generated Human Disease Diagnosis Ontology (HDDO) to identify possible diagnosis from the user's symptom-based queries and PHR data, and to store the user's diagnostic results log. Used inference rule to filter out diseases irrelevant to users and to infer possible diagnosis from the HDDO. Used the TrOWL reasoner to perform semantic inference.	Enhance diagnostic accuracy over existing symptom checker by filtering out irrelevant diseases via semantic inference of the HDDO.
[164]	Integration issue between PHR and hospital information systems.	**SW language:** OWL **Query language:**SPARQL **SW framework :** Apache Jena	Used ontology to semantically represent the rational medical records in hospitals' database, before being stored in the PHR knowledge graph. Used Jena framework as knowledge engine to reason with the OWL data and to gather patients' semantic information.	Enhance data integration between the PHR system and other hospital information systems to realize individual healthcare management.
[165]	Build a modular and interoperable system that can provide IoT services by detecting depression symptoms and recommending services to deliver depressive disorder assistance.	**SW language:** RDF **Query language:**SPARQL **SW framework :** Apache Jena	Used Ontology to semantically represent the virtual objects and link them with sensors and processing resources. Used SPARQL endpoints to extract the customer profile and behavioral data.	Enhance system interoperability, which helps accurately and dynamically select the recommended services for a user, based on his current symptoms and his previous situation.

[166]	1-Integrate Clinical pathways (CPs) with Health Information Systems (HIS) 2-Automate CPs to unstandardized data.	**Ontology editor:** Protégé	Used Ontology to model and define the semantic of CP domain knowledge.	Enhance semantic interoperability among e-CPs and HIS through the CP ontology to communicate shared understanding between heterogeneous applications.
[167]	Aggregate and integrate medical data required to assess patient care quality and guide clinical trials.	**SW language:** RDF **Query language:**SPARQL	Used RDF to represent structured and unstructured data. Used SPARQL query to infer hidden relationships from a wide variety of data.	Enable semantic querying and intelligent retrieval of data in research-oriented scenarios.

Most challenges addressed in the selected papers are mainly related to information management and lack of interoperability between healthcare systems. These challenges hamper the interpretation and analysis of data delaying patients' treatments. Considering all these challenges, researchers integrated the Semantic Web technologies into the healthcare ecosystem to:

- Provide semantic interoperability between healthcare information systems;
- Facilitate the integration of massive heterogenous medical data;
- Enhance data processing (more accurate);
- Enhance medical data analysis;
- Represent various types of health-related data into usable and actionable knowledge.

To understand the Semantic Web's role in enhancing the management of healthcare data, we need to understand its pillars which consist of the standards that enable people to publish, query, reason, track and trace back data on the web. One of the basic building blocks of the Semantic Web is the Resource Description Framework (RDF). RDF is the standard model for web-based data exchange [168]. It is a language that allows anyone to describe, represent and link resources on the web. It consists of a collection of triples: subject, predicate, and object, where the predicate denotes a relationship between the subject and object. Once data are published, we need to query it to retrieve all the results that interest us. SPARQL Protocol And RDF Query Language (SPARQL) is a query language widely used to query data stored in RDF format or ontology languages such as RDFS and Web Ontology Language (OWL). Many authors used SPARQL to improve the search and discovery process needed for a healthcare service. Subbulakshmi S. *et al.* used SPARQL queries to enable dynamic allocation of hospital-based services to meet the users' state of emergency [154]. SPARQL queries are generated based on the users' contextual information, such as location, symptoms, case severity, to retrieve the required services from the hospital ontology OWL file.

The proposed ontology-based semantic retrieval eliminated the elaborate search process thereby helping users not only save time but their lives as well, in situation of emergency.

Another important pillar of the Semantic web is the ontologies that describe knowledge in a way to improve machine automated reasoning. Ontology is a knowledge model that defines a set of concepts and the relationship between those concepts within a domain of interest, rendering them machine-readable and understandable [169]. The two main standards used to define relationships among concepts, write and exchange ontologies/schemas are RDFS [170] and OWL [171]. The choice of the standard depends mainly on the complexity and expressiveness required by a specific application [172]. RDFS is used to describe light ontologies. It offers a range of vocabulary to describe classes of resources and properties that can be maintained by these resources. From that, we can make inferences to discover new relationships from existing knowledge. Since it is built on top of RDFS, OWL represents a richer vocabulary of properties and classes with more logical constructors [171], [173]. Most researchers used the OWL format to represent health-related ontologies because OWL enables automated reasoning capabilities due to the underlying logic, which will allow the inference of further knowledge not represented explicitly within the proposed ontology.
Some authors enhanced data integration between different hospital information systems by using an ontology to:

- Communicate a shared understanding of the structure of information between heterogeneous applications [166].
- Semantically represent medical records to improve patients' data visualizations, comparison, integration [167], and analysis [163], [164].

In a smart healthcare ecosystem, the lack of interoperability among medical devices as well as the lack of interoperability between these devices and the health information system hamper effective data integration in the IoMT. Multiple scholars built and used their ontology to represent the semantic interoperability of the connected medical objects and their data. This ontology not only shares the knowledge in an understandable manner but also becomes a key element in the reasoning task [158], [162]. Using ontology to extract real-time data help clinicians in lifesaving decisions and effective interventions [153].

In addition, some authors introduced the Semantic Web Rule Language (SWRL) into the healthcare ecosystem as a tool to manage and interpret collected data to improve the decision-making of doctors. SWRL is a language used to express rules in terms of OWL concepts and intended for rule-based reasoning. Compared to OWL, SWRL provides more powerful deductive reasoning capabilities. In fact, rule-based reasoning enables the insertion of user-defined rules and the inference of new knowledge [174]. Rhayem A. *et al.* developed several SWRL rules to achieve different goals, notably: help doctors diagnose and provide the appropriate treatment for the patient with hypertension disease based on the detected vital signs, and notify either the patient of the treatment or the doctor for immediate intervention [158]. Sabra S. *et al.* used semantic rules to translate the relationship between risk factors, thereby improving the extraction process of risk factors from the clinical narratives needed for the clinical decision support system (CDSS) to make a diagnosis [157].

To overcome the interoperability problem among the pervasive environment or between different sources of information, some authors used the semantic annotation process to create semantic data that machines can understand and reason upon. Semantic annotation is the process of linking electronic resources to a specific ontology [175]. Karthik N. *et al.* used semantic sensor network (SSN) ontology to unify the representation of sensor data where data can be shared and reused among various pervasive applications [156]. Additionally, Dridi A. *et al.* proposed a flexible semantic annotation model for data interpretation that annotates data according to their sources: SMD ontology for the connected medical devices, and UMLS ontology for the medical documents [161]. Semantic annotation helps data integration, data exchange, data reuse, and information discovery [175]. And, to enhance the diagnostic accuracy of disease and facilitate the intelligent detection of emergency cases, they integrated to their semantic module a semantic reasoner to insure the interpretation of the medical data. Besides the above mentioned ontologies, several others have been created to describe a specific domain in biomedicine, notably the SNOMED Clinical Terms (SNOMED-CT) that represent the clinical concepts, and the Foundation Model of Anatomy (FMA) that describe human anatomy [176].

Moreover, some authors introduced ontology mapping to deal with the challenge of managing multiple ontologies. Ontology mapping determines correspondences among the concepts belonging to separate source ontologies. It establishes the linkage between individual domains, and it enables reasoning as users can query different local ontologies through an integrated global ontology [177]. Sigwele T. *et al.* used ontology mapping to exchange knowledge seamlessly and to enable the collaboration between distinct

healthcare systems with different underlying ontologies [155].

In the next section, we will describe our solution design where we combine blockchain and semantic technologies to limit the adverse drug reactions due to the fragmented care in the Lebanese healthcare system. Since data fragmentation is a serious issue that affects and delays patients' treatment, we consider that dealing with the proposed use case could be the first step towards a more global solution for improving the healthcare system.

4.4 Design overview

To limit adverse drug reactions due to inappropriate medicine intake, we propose to use a blockchain-based system to store and share the patients' prescriptions and drug allergies with their physicians. The blockchain provides integrity, accessibility, reliability, and traceability of the data for users. Having the history of patients' drug allergies and medications constantly accessible can help physicians save time when providing the appropriate care, especially in the case of an emergency.

Based on these stored data, the system would alert the physician of any possible drug reaction (allergy or interaction effect) when an unsuitable medicine is being prescribed. On the other hand, knowing that the semantic description of the drugs enables inference and knowledge discovery, we propose to store the semantic description of the drugs and their composition, as well as all relevant information, in a graph database fed by an ontology. For this purpose, we defined a drug ontology that includes the classes of drugs, active pharmaceutical ingredients, drug allergies, and the relationships among them required in our use case to verify the following statements:

- The patient is not allergic to the active substance of the newly prescribed drug.
- The newly prescribed drug has no interaction with any of the patients' current medications.
- The newly prescribed drug is not similar to any of the patients' current medications.

Once the previous statements are validated, we can store the newly prescribed drug on the blockchain.

Practically speaking, it means the patient can start taking the drugs safely after acquiring them from the pharmacy.
We identified in our framework four actors and their respective roles:

- The physician who is responsible for prescribing drugs to patients and adding allergy incidents to their records.
- The patient who is responsible for controlling the access to their medical record by sharing it with the appropriate users.
- The administrator who is responsible for creating or updating participants demographics (i.e., patients, physicians, auditors)
- The auditor who is responsible for tracing back all the events/transactions that occurred on the blockchain network. In our case, the auditor's role is limited to obtaining audit evidence directly from the blockchain, especially in the case of a complaint. However, in future work, the auditor role will evolve and include procedures that assess the prescription quality as defined by the World Health Organization (WHO). The WHO created in collaboration with the International Network for Rational Use of Drugs a set of "core drug use indicators" to study the rational use of drugs in developing countries [178].

Figure 4.1 shows the UML use case diagram with the main four actors and the main use case scenarios. We will discuss all related transactions in the next subsection. It is important to note, that the drug ontology is stored off-chain on a graph database. This ontology will be accessed each time a transaction needs to be added to the blockchain after the assessment and validation of the required statements as previously explained.

4.4.1 Blockchain selection and framework

Since our proposed solution must be HIPAA compliant, we must ensure that only authorized entities can view or update the data. Hence, we chose to adopt a permissioned blockchain: the Hyperledger Fabric. In a permissioned blockchain, a centralized or decentralized authority controls who can issue

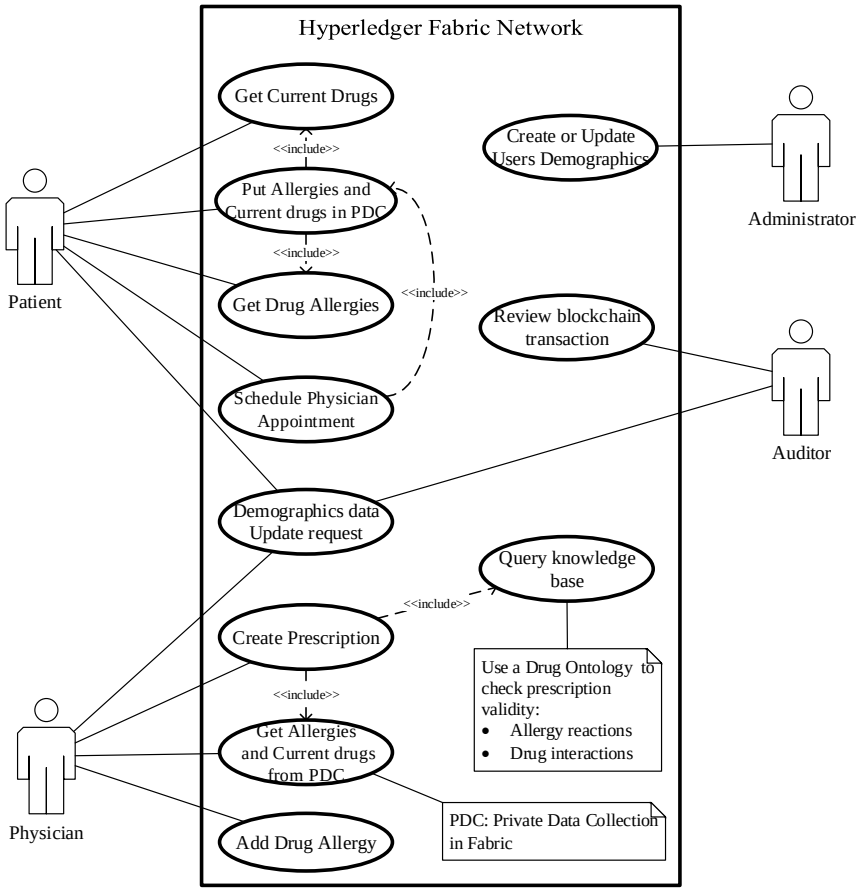


Figure 4.1: UML use case diagram illustrating the interaction between the different system actors

transactions, and all participants are known to each other [28]. Hyperledger Fabric is, in fact, one of the most mature technology projects in the Hyperledger platform, as well as the most popular due to its flexibility, modularity, and rich documentation. It allows pluggable consensus protocols, notably Crash Fault Tolerance (CFT) and Byzantine Fault Tolerance (BFT). The latter is still under development and will later replace Raft, a CFT ordering service in Hyperledger Fabric [45]. In addition, Hyperledger Fabric offers developers multiple programming languages to write smart contracts (aka chaincode), such as JavaScript, TypeScript, Golang, Java, and Node.js [45]. Compared to the permissionless blockchain, such as Bitcoin and Ethereum, Hyperledger handles only business logic through the chaincode functionality and enforces privacy and confidentiality through its channel architecture and private data [179].

As depicted in Figure 4.2, three organizations are collaborating in the proposed healthcare ecosystem, notably the healthcare authority, the health information systems auditor, and the Lebanese Order of Physicians. Each of these three organizations has at least two peers, an orderer, and a membership

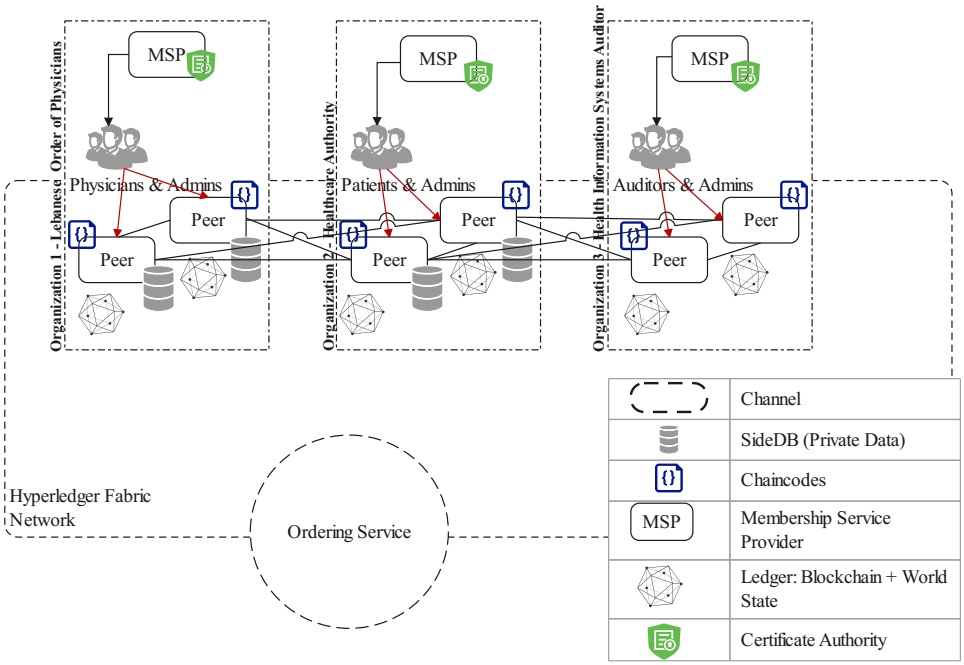


Figure 4.2: Hyperledger Fabric-based network topology for management of medication histories and drug allergies of patients

service provider (MSP) with their certificate authority (CA). The MSP enables identity validation and signature verification by and for all channel members, a key element in the network [45]. Even though it is possible to have one MSP serving multiple organizations, it is not recommended since it compromises data privacy [45]. In that case, the organization-scoped messages will propagate not only to the peers belonging to the same organization but to all the network peers since all peers' identities are under the same MSP. Therefore, each organization will use its certificate authority to generate identities for its network participants/members. These identities are required to interact with the Fabric network. The MSP will manage these identities by assigning them roles and permissions. We identified four different roles [45], [179]:

- **Client:** it represents applications, end users or any identity that interacts with a peer to communicate with the blockchain by invoking transactions in the network.
- **Peer:** it represents any identity that endorses and commits transactions. There are two types of peers: the endorsing and the committing peers. The endorser verifies the client signature and executes a chaincode function to simulate the transaction. The committing peer verifies the endorsements and validates the transaction results before adding the transactions into the blockchain.

- **Admin:** it represents any identity that handles administrative tasks such as creating channels, running peers and helping them join the network, approving and installing chaincode, changing block configuration, executing ordering service, etc.
- **Orderer:** it represents any identity responsible for ordering the transactions and assembling them into blocks that peers will append to the blockchain. In the designed network, the ordering service consists of three orderer nodes, each belonging to one of the three organizations. We decided to represent them as a separate entity, as shown in Figure 4.2, to differentiate them from the other peers and highlight the key role of the ordering service in the Hyperledger Fabric network.

Peers and orderers work together to manage consensus, required to [45], [179]:

- Achieve agreement on the different transactions within a block in terms of order and correctness;
- Create trust and security across the decentralized network;
- Keep synchronized the multiple copies of the ledger in the network.

Consensus is achieved through three steps: endorsement, ordering, and validation. The ordering service can be implemented as a centralized or decentralized service. In our design solution, we propose to adopt Raft, a CFT ordering service, because it allows every organization to have its ordering node hence, achieving a level of decentralization not available in the other ordering services (notably Kafka and Solo) [45].

In terms of privacy and access control, we must ensure that only certain authorized entities and, for a limited period, access patient health information. To meet HIPAA rules, we choose to give patients control over their data by giving them the responsibility of sharing them with the consultant physician of their choice.
As a permissioned blockchain platform, Hyperledger Fabric has several constructs that help achieve privacy. Two of these constructs are the channel and the private data collections. A channel is a private "subnet" of communication between two or more network members enabling them to transact privately. Each channel defines a single ledger for all its transactions and state changes [45], [179], [180]. In our concept design, we deploy one channel where organizations will collaborate by executing transactions. Each peer in the

network hosts a copy of the ledger and the chaincode. Chaincode is a piece of code that encapsulates the business logic that, once executed, may modify the ledger [45], [179], [180]. The ledger contains the blockchain, which stores the history of all transactions on a particular channel, and the state database (aka world state), which represents the current value of all assets in the ledger [45]. The state database supports LevelDB and CouchDB. We will be using the CouchDB since it permits the deployment of indexes within a chaincode to make queries more efficient and to enable large datasets querying. Compared to LevelDB, CouchDB fulfills chaincode auditing and simple reporting criteria. It is recommended to store data in JSON format to leverage the benefits of CouchDB [45].

Since all channel members have access to the data transacted on the blockchain, we need to find a solution to share data between a subnet of participants within a channel without creating a new one. Hence, we introduced private data collection (aka SideDB) into our ecosystem to prevent the administrative overhead introduced by the creation of multiple channels [45]. These collection data will be shared between authorized peers through gossip protocol, per the collection policy, and the blockchain will only retain the private data hashes [180]. Moreover, collection data is kept secret from the ordering service, which is advantageous, especially if the ordering service is administered by an organization that is not permitted to see the data. Furthermore, there is a collection property called blockToLive that indicates in terms of blocks how long the data should remain on the private database before being purged. Hence, after a certain number of created blocks, the data will no more be available. The data hashes, already committed to the blockchain, will be used as proof of transactions for audit purposes [45]. As shown in Figure 4.2, the private data between the Lebanese Order of Physicians and the healthcare authority is managed by their respective peers. However, in the future, we plan to expand our ecosystem to include pharmacists and health insurers; we will then build additional private data collections to help patients share their data with other system actors. For each private data collection created, we define mainly the organization peers allowed to store private data, the retention period of the data in the SideDB in terms of blocks, and the endorsement policy that should be satisfied to write to the private data collection [45]. Even though we can enforce the read or write access to private data by setting the collection fields "memberOnlyWrite" and "memberOnlyRead" to true, we will be setting them to false to apply a more granular access control in the chaincode logic.

We will make access control decisions based on the client certificate. Hence, the client will invoke a chaincode according to the attributes and the Organizational Unit associated with his identity. In our use case scenario, the access control rules of the main network transactions are implemented as follows:

- Any user having "Patient" as attribute and belonging to the "healthcare authority" organization can:
  - query the ledger to get his currently prescribed drugs through the *GetCurrentDrug* function, which takes the patient ID as a parameter. To identify the current patient's medications, we will check the validity of the patient's prescriptions for the past twelve months starting from the current date. It consists of controlling the duration of treatment of all these prescriptions. According to the dispensing guidelines for pharmacists in Lebanon, the maximum validity of a drug prescription for acute conditions is one month after the date of issue and one year for chronic disease [181].
  - get the list of drugs that induced allergic reaction through the *GetDrugAllergies* function, which takes the patient ID as a parameter. The list of drug allergies is retrieved directly from the world state.
  - write his current drugs and drug allergies in the private data collection shared between the healthcare authority and Lebanese Order of Physicians organizations through the *PutPrivateDrugsAllergies* function, which takes the patient ID, physician ID and collection name as parameters. This function includes the *GetCurrentDrug* and *GetDrugAllergies*, which will be executed once the *PutPrivateDrugsAllergies* function is invoked.
  - request an appointment from a physician through the *ScheduleAppointmentPhys* function, which takes the patient and physician ID as parameters. This function includes the *PutPrivateDrugsAllergies* function, which will be executed once the *ScheduleAppointmentPhys* function is invoked.
- Any user having "Physician" as attribute and belonging to the "Lebanese Order of Physicians" organization can:
  - read from the private data collection shared between the healthcare authority and Lebanese Order of Physicians organizations through the *GetPrivateDrugsAllergies* function, which takes the patient ID, physician ID and the collection name as parameters. The physician will be able to retrieve the patient's current medication and drug allergies required for validating the new prescription.
  - write the patient prescription into the ledger through the *CreatePrescription* function, which takes the physician ID, patient ID, prescription date, and the list of drugs prescribed with their required details as parameters. This function includes two other functions: the *GetPri-*

vateDrugsAllergies and a function that queries the drug ontology to validate the prescription (no drug allergy or drug interactions detected). When adding a valid prescription to the ledger, we will add the duration of treatment, the drug dosage, and the proprietary name of the drug. We will represent each drug with a unique URI. It is required since we are pointing to the remote graph database that integrates the knowledge graph of our drug data through the chaincode. We will query our drug ontology to retrieve all the knowledge needed to validate all our statements. For example, a drug will have this type of URI: `http://www.semanticweb.org/healthcare/ontology#Xanax`.

- write the patient allergy into the ledger through the *AddDrugAllergy* function, which takes the physician ID, patient ID, and the drug that induced an allergy reaction with all its required details as parameters. As per the National Institute for Health and Care Excellence (NICE), some information should be documented when a patient presents a suspected drug allergy, notably: the generic and proprietary name of the drug, a description of the reaction, the number of doses taken before the reaction occurred, the drug classes to avoid in future, and the drug indication [182]. In our case, when adding a drug allergy to the ledger, we will be pointing to the active pharmaceutical ingredient and the proprietary name of the drug located in the remote graph database, both required to minimize the patient's exposure to a similar drug. In this way, we are representing the drug and its main component with a unique URI. Also, we will include a pointer to the drug allergy indication in the graph database to give a brief description of the reaction.

- Any user having "Patient Administrator" as attribute and belonging to the "healthcare authority" organization can:

  - write the patient demographics into the ledger through the *CreatePatient* function, which takes the patient: ID, name, address, date of birth, sex, address, phone number, and contact number in case of emergency as parameters.

    There is an administrator for each organization. He will be adding the demographics of the patients, auditors, and physicians, respectively. But in the UML use case diagram (Figure 4.1), we represented the administrator as a single entity that manages the different system actors to avoid cluttering the diagram. Each function will have its input parameters. For example, for the physicians, we will list the hospitals or clinics they work at in addition to their field of

expertise in medicine. Besides, we need to differentiate the user having “administrator” as an attribute, responsible for creating system actors, from the identity registered as “admin” role that conducts administrative activity at the MSP, peer, or organization level.

- Any user having “Auditor” as an attribute and belonging to the "health information systems auditor" organization can:
  - query the transaction information history on the fabric blockchain network to increase accountability and validate transaction integrity from a legal, audit, or compliance perspective.

Furthermore, to maintain data privacy and guarantee who can get to see the data, we chose to encrypt patient data and user demographic data, each with their respective public key. We will introduce application-level encryption, having the administrators and physicians encrypt the transaction input parameters with the required public key, except the patient ID, physician ID, and prescription date, to maintain searchability. However, when invoking the *CreatePrescription* function, the encryption must be done after the prescription validation. In this case, the encryption key will be passed as transient data to maintain its confidentiality because transient data doesn’t stay in the transaction record. As for the *PutPrivateDrugsAllergies* function, the patient will have to encrypt the data with the physician’s public key before writing it to the private data collection.

4.4.2 Drug ontology design and formalization

To access vital domain knowledge and assert domain inference rules for the purpose of reducing inappropriate medication intake, we defined a drug ontology that will be queried using the SPARQL query language each time a prescription transaction is added to the blockchain. The drug ontology includes the classes of drugs, active pharmaceutical ingredients, allergy reactions, and the relationships between them in order to notify the physician of any possible drug allergic reaction or drug interaction. We chose OWL as a semantic language to describe our ontology since it provides rich semantics and expressiveness and helps us derive implicit associations between different entities through a reasoner module. OWL supports data reasoning that allows developers to reduce the data stored explicitly and reduce the queries used to retrieve

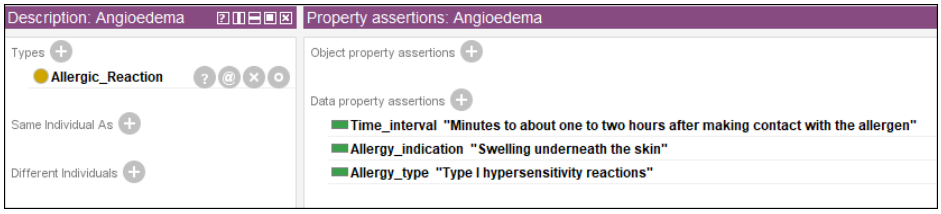


Figure 4.3: Example of datatype properties of an allergic reaction: Angioedema

those data. Compared to RDFS, OWL offers more expressive class definitions (such as class intersection, class union, etc.) and more expressive property definitions, notably the object properties and datatype properties [171].

A datatype property links an object with a data value (Figure 4.3). Some of the data properties used in our projects are:

- Maximum Daily Exposure (MDE)
- Allergy indication: it represents the signs and symptoms of the allergic reaction
- Anatomical Therapeutic Chemical Code (ATC code): it represents a unique code assigned to a drug according to its properties and the organ or system on which they act.
- Chemical Abstracts Service Registry Number (CAS RN): it is a unique identifier assigned to every chemical substance.

An object property determines the relationship between two objects, hence enabling the inference process (Figure 4.4). It's the reasoner that will oversee the inference process, which involves an automated discovery of new facts based on given data.

The object properties used in our projects are:

- **Drug-Drug Relationship:** we address here the **"HasInteractionEffect"** object property that will automatically search for any incompatibility between prescribed drugs. The main interaction types are duplication, opposition and alteration. A drug-drug interaction can either increase or decrease the drug effect. By decreasing it, it may cancel the desired effect and by increasing it, it may lead to an inappropriate result.
- **Drug-Main Component Relationship:** we address here the **"HasActiveSubstance"** property that

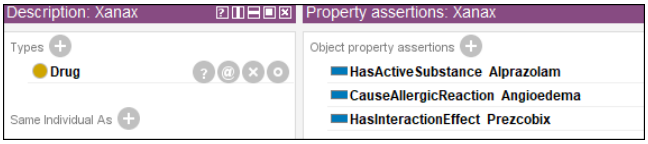


Figure 4.4: Example of object properties of a drug: Xanax

will automatically search for similar components between two drugs. A doctor, who may not be familiar with some prescribed drugs, may give a similar prescription to the patients hence increasing the dosage of an active ingredient.

- **Drug-Allergy Relationship:** we address here the **"CauseAllergicReaction"** property that will automatically inform the patients and physicians of possible allergic reactions related to the prescribed drug.

Taking into consideration the previous properties, we managed to define our drug ontology concepts, as shown in Figure 4.5, which covers all the necessary relationships for dealing with the drug prescription issue. We chose to store our ontology in a graph database to reduce the query response time. Elchamaa R. *et al.* [183] and Lampoltshammer T. *et al.* [184] highlighted in their paper the advantage of a graph database approach and proved it more appropriate for real-time applications than pure ontology-based approaches and reasoning. By introducing a graph database environment, Lampoltshammer T. *et al.* [184] improved the ontology-based classification of segmented remote sensing data; the results showed a reduction of

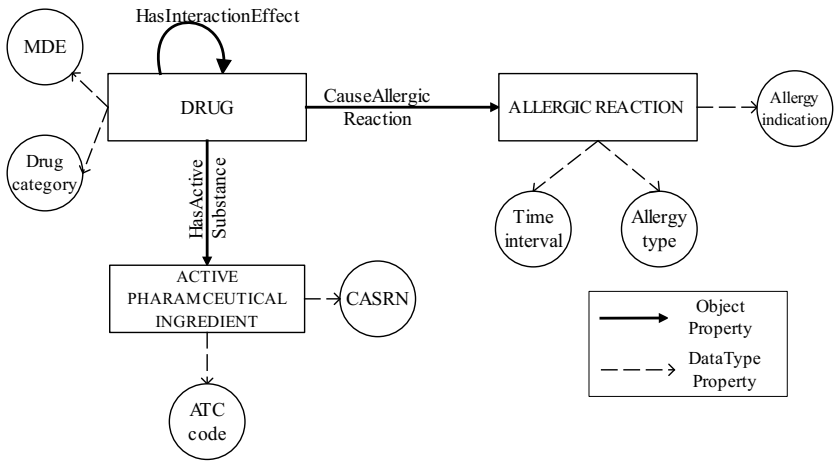


Figure 4.5: Drug ontology concept

the computational time for the classification tasks while maintaining the same level of accuracy. Graph Database Management Systems (GDBMS) are NoSQL databases that use graph structures to store and represent data. The graph consists of nodes that represent data entities and edges that represent relationships between entities. Graph databases enhance the visualization of these relationships, making them useful for heavily interconnected data. It is not the case in a relational database where concepts are stored according to a defined schema with no information about their meaning and relationships to other concepts [183]. Besides, Graph databases, unlike relational databases, do not contain redundant data, are more flexible when adding a new relationship, and do not require complex queries when dealing with inter-related datasets [184], [185]. To select the most suitable graph database, we checked the ranking of GDBMS and RDF stores according to DB-Engines [186], [187]. Since we are interested in GDBMS that supports RDF triples and SPARQL queries, we identified Virtuoso [188] and GraphDB [29] in both lists, with positive community feedback and available online resources.

As a reasoning strategy, we will adopt the forward-chaining since the query answering can be completed promptly once all inferences have been computed [183], [189], [190]. Forward chaining is a data-driven inference technique that infers new facts from known facts. On the other hand, backward reasoning is a goal-driven inference that can be computationally expensive, principally when dealing with complex graphs, because it involves reasoning every time a query is answered [189], [190]. The goal of forward reasoning, in our case, is to answer queries that consist of searching and retrieving from the ontology all drugs that interact with the newly prescribed medicine or that share a similar active substance. Virtuoso does not support forward chaining; hence, we will be uploading our drug ontology to a GraphDB repository.

In order, to assess the three statements that we have put forward, we need to query the GraphDB as follows:

- Select from the GraphDB all the drugs that **HasInteractionEffect** with the drug the patient is taking. This query takes the patient's current drug and newly prescribed drug as a parameter to alert us of possible drug-drug interactions.
- Select from the GraphDB the active pharmaceutical ingredient related to the drug the patient is taking and the newly prescribed drug (**HasActiveSubstance**). This query takes the patient's current drug and newly prescribed drug as parameters to alert us of a similar drug prescription.

- Select from the GraphDB the active pharmaceutical ingredient of newly prescribed drug (**HasActiveSubstance**). This query takes the active ingredient to which the patient is allergic and the patient's newly prescribed drug as parameters to alert us of possible allergic reaction.

4.4.3 System interaction

In this subsection, we will discuss the interaction of the network users with the Hyperldger Fabric blockchain, and the drug ontology stored in GraphDB.

Figure 4.6 shows the UML sequence diagram of a use case scenario, having the patient requesting an appointment from a specific physician and the physician submitting a prescription following the patient appointment. Note that we condensed the sequence diagram to include only the main components needed to understand how our proposed solution operates. A user, whether a patient or physician, connects to the blockchain network through a web interface. To participate in the Hyperledger Fabric network, each user should have a certificate and a private key. These keys and certificates are issued by the Fabric Certificate Authority (Fabric CA). Once the user has the enrolment certificate, he can submit a transaction via the web application. Therefore, a PHP script is invoked to trigger the execution of the appropriate chaincode function. The endorsing peers will verify the transaction legitimacy by verifying the user identity and authorization to perform the proposed operation on that channel. Since the user is a "Patient" and belongs to the "healthcare authority" organization, the endorsing peers execute the chaincode. The execution of the *ScheduleAppointmentPhys* function, which takes the patient ID and Doctor ID as parameters, will trigger the *PutPrivateDrugsAllergies* function. This function consists of encrypting the patient allergies and current drugs by the doctor public key before writing them in the private data collection.

Hence, the simulation result, which is the actual private data, will be stored on the SideDB of the Lebanese Order of Physicians and the healthcare authority organizations' peers. The hash of these data will be endorsed, ordered, and written to the ledger. An "alert" pop-up is displayed to notify the patient of the success or failure of the transaction. Now, it is up to the physician to submit a transaction. Endorsing peers will execute the chaincode since the physician has the right to create a prescription. The *CreatePrescription* function consists of getting the patient's allergies and current drugs from the private data collection. The physician decrypts these data using his private key before passing them as parameters to the *CheckPrescriptionValid* function alongside the newly prescribed drugs. The *CheckPrescriptionValid* function will

query the GraphDB that stores the knowledge graph related to our data. It will fetch for any similarities or possible interactions between the patient's current drugs and the new prescription to avoid any drug overdose, drug interaction, or potential allergy reaction. Once the prescription is validated, it is encrypted by the patient public key and added to the ledger. An "alert" pop-up is displayed to notify the physician of the success or failure of the transaction.

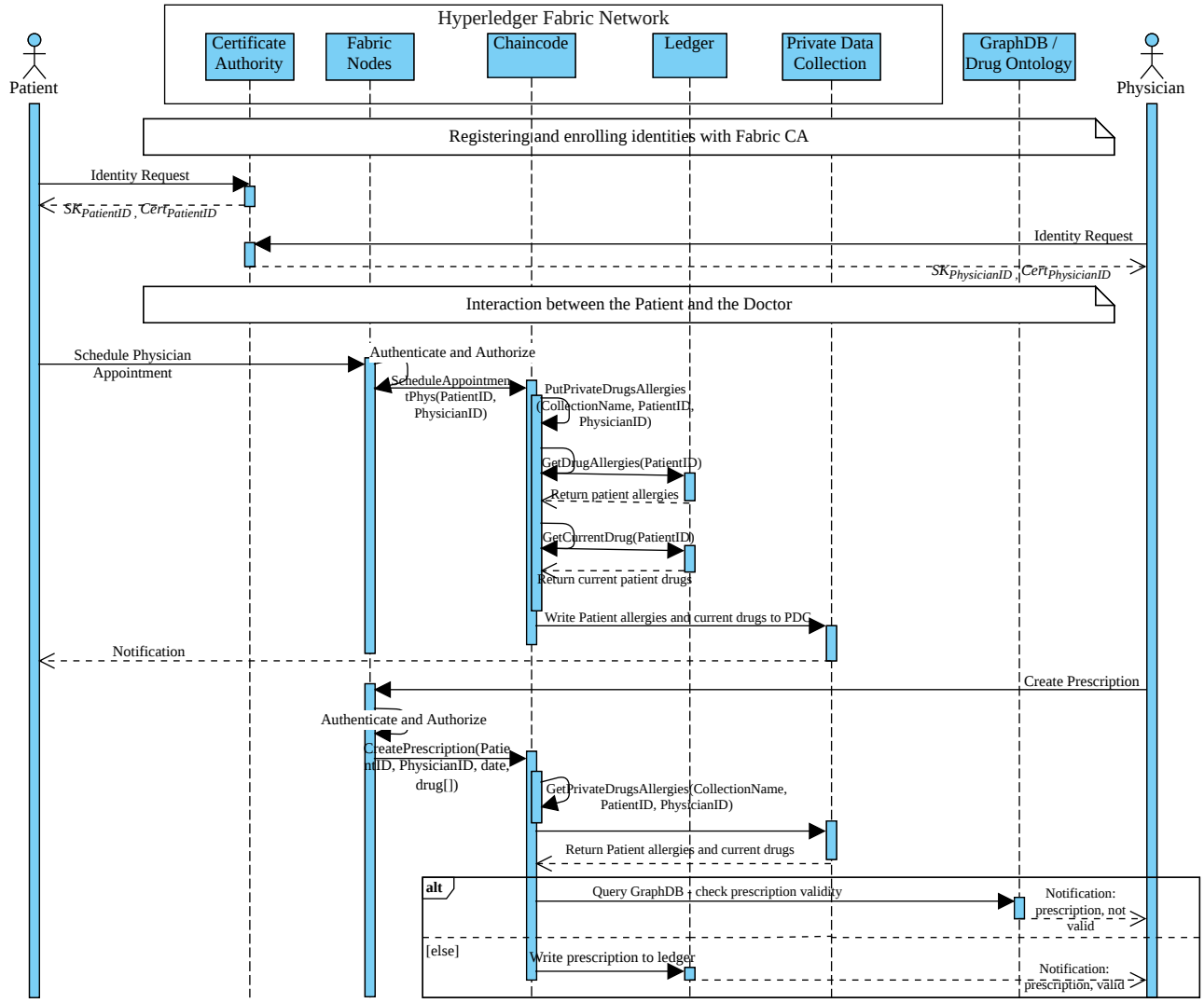


Figure 4.6: UML Sequence Diagram of a use case scenario between the patient and physician

4.5 Discussion

Our proposed framework enhances the management of patients' data, notably their drug prescriptions and drug allergies. It enhances the physician-patient interaction by enabling effective communication through the deployment of Hyperledger Fabric blockchain and drug ontology. The blockchain is used to share patients' data securely while allowing them to control the access permission since the blockchain provides availability, integrity, immutability, and transparency of shared data.

The drug ontology which is stored off-chain, is used to limit adverse drug reactions that might result from an inadequate prescription. SPARQL queries are used to retrieve all the information needed from the ontology in order to notify the physician of a potential adverse reaction, such as an allergic reaction to the newly prescribed drug or possible drug interaction with his current medication. We will use the chaincode to connect to the GraphDB and query our ontology. Upon verifying that there is no risk for the patient, the transaction will be validated and appended to the blockchain.

Besides integrating the drug-allergy alerts and drug-drug interaction alerts to enhance the treatment process, our proposed system ensures patient privacy and is HIPAA compliant since patients have rights over their data. In fact, there are many controls that brought security and privacy to our ecosystem, we cite:

- Access Control: provided through the implementation of identity-based and organizational-unit-based access policies into the chaincode logic to manage access to resources. However, access control is also provided by membership service providers, who manage users' identities, authenticate all network users, and ensure that only members with valid memberships can transact on the blockchain.

- Data encryption: provided through the application-level encryption in order to maintain network participants' privacy. All data is encrypted whether, stored on the ledger or on the private data collection. Only the user with the correspondent private key can decrypt it.

- Private data collection: introduced as a tool to help patients gain control over their data by deciding what and with whom to share their private information.

In addition, accountability is provided since all transactions are signed and cannot be forged, hence no one can deny responsibility towards their actions on the designed network.

Even though our proposed system enhances the treatment process, we still have to study the efficiency of our solution in terms of throughput and latency, especially after introducing the encryption process into our ecosystem. In addition, we need to choose the adequate value of the blockToLive parameter that gives the physician enough time to check the data before it becomes unavailable on the network. Moreover, patients are not necessarily allergic to the drug's active ingredients only but also to a particular excipient, an additive substance used in drug and vaccine manufacturing. Hence, we need to enhance our ontology to take into account excipients and more extensively vaccines.

4.6 Conclusion

Healthcare data management can be very challenging, specifically when dealing with fragmented data and centralized data stores, both of which hinder data sharing and jeopardize patients' privacy. In this chapter, we addressed the health data fragmentation issues by using a case study approach to examine a challenge in the Lebanese healthcare system at the level of drug prescription: the adverse drug interactions/reactions. We proposed to combine blockchain technology with Semantic technologies as a solution to help patients manage their health data. We used the blockchain to store and share the drugs' prescriptions while allowing the patients to manage the access permission to their personal data. Moreover, we defined a drug ontology that includes the classes of drugs, active pharmaceutical ingredients, drug allergies, and the relationships between them (**HasInteractionEffect**, **CauseAllergicReaction**, **HasActiveSusbtance**) required to notify the physician of any possible drug allergic reaction or drug interaction. Our findings showed that the semantic description of data enhanced the blockchain-based healthcare ecosystem because they enabled knowledge inference from a few given facts, notably the patients' allergies and current drug prescriptions. Furthermore, accurate and traceable recording of patients' data improved patients' safety.
The adopted use case showed how the blockchain and the semantic data representation contributed to healthcare data management by addressing the data fragmentation issue. With its immutability and transparency features, the blockchain provided a secure infrastructure to handle data while reducing the

risk of unauthorized access and by preventing data tampering through a tamper-proof audit log. The ontology served as an intelligent representation of information resources, and it was queried from a graph database, which enabled fast automated reasoning and interpretation of data.

In the future, we will develop our ontology to include more information regarding medication safety and contraindications (notably data related to vaccines), which could enhance our proposed system and address the challenges of any healthcare system at a wider scope. We intend to expand our design to include pharmacists and insurance providers before integrating it into EHR systems.

Chapter 5

Supply Chain meets Care Chain: Navigating Healthcare Privacy Challenges

5.1 Introduction

Health data are undeniably the most valuable asset retained by healthcare providers due to their significance in advancing clinical research and improving healthcare outcomes. This importance has led to a growing demand from pharmaceutical companies for increased access to accelerate the discovery and development process. Leveraging data from electronic medical records (EMRs), clinical trials, and medical sensor devices allows pharmaceutical companies to target specific populations and enhance treatment effectiveness [191], [192].

Conversely, it is crucial to provide patients with the ability to access, download, and share their health data with individuals or entities of their choice. When patients can easily access their health data, they become more actively involved in their healthcare journey, enabling them to comprehend their health conditions, track their progress, and make well-informed decisions regarding their treatment. For example, patients can use their health data to participate in research or clinical trials. Furthermore, giving patients control over their health data cultivates trust between patients and healthcare providers, fostering a more

collaborative and patient-centered approach to care.

However, patients currently lack control over the ways in which pharmaceutical companies access their data, as well as how hospitals and clinics manage and share their information [193]. In fact, health data are sold by electronic medical record companies for the benefit of health data mining companies, which will turn raw data into actionable information sold to interested buyers, notably pharmaceutical companies and insurers [31]–[34]. For instance, Phreesia, a software company, provides healthcare organizations with applications to streamline patient intake processes. While offering its software solutions, Phreesia also engages in selling advertisements to pharmaceutical companies. These ads are tailored based on the information entered by users according to their medical needs [194]. Moreover, pharmacies, health insurers, and hospitals receive money for sharing details about patients' health conditions or prescribed medications [31]. Although healthcare data are theoretically stripped of personally identifying information before being shared with data collectors, some healthcare organizations which are covered by the Health Insurance Portability and Accountability Act (HIPAA) have been recently accused of sharing identifiable health data with social media networks [31], [35]. Making health records available is essential to provide patients with the care they need. However, when health information is shared without patients' consent encourages them to conceal information due to a lack of confidence in the security of the healthcare information system [195]. Additionally, when patients need to access or get their medical records, healthcare providers meet their requests with delays and fees while being selective about the shared information. Health record retention by health providers make them in full control of information with high commercial and research value [36]. Furthermore, healthcare organizations have become the main target of hacking attacks since health data are more appealing than credit card information [37]. Among recent incidents, we cite the ransomware attack that hit the André Mignot hospital in Versailles on December 3, 2022. The hospital was forced to shut down its network as a security measure, suspend operations, and transfer some patients to other hospitals [196]. Data breaches occur when hackers or unauthorized parties gain access to sensitive or confidential information without permission. These data breaches affect patients' safety, damage healthcare organizations' reputations, and increase their financial burden. Ensuring data privacy and security is the foremost concern of healthcare organizations [197], [198]. Consequently, these entities should adhere to the privacy requirements set forth by HIPAA.

The new federal rule on interoperability and information blocking compels healthcare organizations

to give patients unfettered access to their health records in an electronic format without delay or charge. And as of October 6, 2022, the rule has been updated, and the definition of electronic health information expanded beyond the United States Core Data for Interoperability (USCDI) Version 1 to include electronic Protected Health Information (ePHI) [199], [200]. It consists notably of the medical and billing records, the enrollment, and claims adjudication except for the psychotherapy notes and information compiled for use in a civil, criminal, or administrative proceeding or action [199], [200]. Besides giving patients the right to obtain a copy of their health records, HIPAA gives them the right to choose with whom to share these records, except when the patient is unable to provide consent when disclosure is necessary for public health, by law, or for emergency treatment [201], [202]. Nevertheless, the claim of promoting public health and safety is sometimes used to obtain patient data under false pretenses [31]. While de-identified patient health information is no longer subject to HIPAA regulations, the advancements in computing technology, coupled with the vast amount of rich medical data collected from applications or smart devices, have significantly aided data scientists in the process of reconstructing individual identities [31], [34], [203]. Therefore, safeguarding anonymized patient data is essential to mitigate the risk of re-identification, especially in cases involving patients with rare health conditions.

Blockchain technology has emerged as a solution to ensure the security of health records [38]–[40]. Blockchain enables decentralized management, provides an immutable audit trail, and ensures transparency of information which adds a degree of accountability to the system [28]. But what about data confidentiality and privacy? Even though actual blockchain technology offers several benefits, such as transparency, reliability, integrity, and availability, some challenges still need to be addressed, notably the privacy issues of blockchain-based healthcare applications [38], [204]–[206].

This work aims to give patients more control over their health data. Healthcare organizations should prioritize the privacy and preferences of patients before handling or sharing their data. The right balance between data utilization and privacy must be carefully established, as a lack of transparency in data handling can hinder effective care coordination and compromise patients' confidentiality. Consequently, we propose a new blockchain-based healthcare solution that addresses all the above challenges, including blockchain limitations. The proposed solution must: (1) enable the seamless exchange of health information among the different system actors, such as patients, physicians, and pharmaceutical company researchers (2) protect patients' data from being exposed to unauthorized parties or tracked and traded without the patients' consent

(3) ensure the privacy and security of patient health information during the entire data life cycle.

The rest of this work is organized as follows. In Section 5.2, we conduct a comprehensive literature review of the privacy and security challenges within the current blockchain-based healthcare ecosystem. We examine various approaches documented in the literature to address these concerns and subsequently identify the key factors to achieve privacy and security in our proposed framework tailored for an enhanced blockchain-based environment. Sections 5.3 and 5.4 describe, respectively, the different components and the implementation details of our proposed healthcare data governance solution. In Section 5.5, we present the security validation and verification of the proposed system using Automated Validation of Internet Security Protocols and Applications (AVISPA). Section 5.6 assesses the system's capacity to ensure secure and private interactions among various healthcare stakeholders. In this section, we additionally provide a roadmap outlining a clear and structured methodology for adopting our proposed framework in any data-driven domain. Finally, in Section 5.7, we present our conclusions, discuss challenges, and outline avenues for future work.

5.2 Literature review on blockchain-based approaches for data protection in the healthcare ecosystem

To identify the most appropriate measures for enhancing the privacy and security of health records within our blockchain-based healthcare solution and address the associated shortcomings, we have conducted an extensive analysis of the literature on blockchain integration within the healthcare system.

5.2.1 Blockchain strengths and weaknesses

A blockchain is a distributed ledger that records and shares all transactions that occur within the blockchain network. It provides a permanent record of transactions ordered into an immutable block [28]. The blockchain has been introduced as a tool to address the vulnerabilities within the centralized design of

current healthcare services. It aims to tackle the security and data access control issues that prevented secure information sharing among peers. Therefore, it is used as a platform for sharing and managing access rights to healthcare records while ensuring a transparent audit trail of shared records [67], [75], [92], [104]. Our thorough investigation of the literature has shown that the blockchain cannot store large datasets due to its scalability issue. The scalability problem is related to the block size [106]. According to Dinh *et al.*, the increase in block size leads to a proportional decrease in block generation rate, thus hindering the overall throughput [43]. For example, in a bitcoin network, the block size should not exceed 4MB to get a maximum throughput of at most 27 transactions/s [107]. Hence, storing the healthcare data on blockchain will impact its performance considering that a blockchain of hundreds of petabytes requires significant computing power and increased network bandwidth [90]. In fact, the majority of the papers we have analyzed address the blockchain limited storage capacity by adopting a combination of on-chain and off-chain storage. Moreover, the immutable nature of the blockchain makes it unsuitable for storing patient data since once data is added, it cannot be deleted. This inability to delete data does not align with the General Data Protection Regulation (GDPR), which requires organizations to be capable of removing personal data upon individuals' request.

We have also found that cryptographic techniques play a crucial role in ensuring the integrity, authenticity, immutability, and non-repudiation of the blockchain ledgers because even an authenticated node can act maliciously. These techniques include: the root hash of the hash tree which is used to detect data tampering and validate transactions; the block header hash which is used to verify the integrity of the block and transactions; the asymmetric cryptography which ensures the non-repudiation and authenticity of transactions and the integrity of the transmitted data in the blockchain [28]. Although the blockchain offers various advantages in terms of security, it also presents vulnerabilities that can compromise the security of shared data and users' privacy, especially when dealing with a permissionless blockchain [206], [207]. In fact, by monitoring the transactions within a permissionless blockchain, there is a risk of revealing the patients' identity through linking sufficient data associated with the patient [206]–[208]. Zhao *et al.*,for example, emphasized the importance of implementing a security mechanism to protect data privacy in permissionless blockchain before introducing it into the healthcare ecosystem [209]. Furthermore, inherent cybersecurity risks are not eliminated by blockchain technology, and these risks are often associated with human error or intentional actions that can lead to vulnerabilities that cybercriminals can exploit [28]. In addition, the decentralized nature of blockchains means that data are stored and shared among many

nodes, making it difficult to control those with access to information. Malicious actors can take advantage of weaknesses in blockchain networks and launch attacks like zero-day attacks. Blockchain platforms and smart contracts can also be targets for denial-of-service attacks [28]. In view of these risks, and as hackers become more familiar with blockchain networks and their vulnerabilities, it becomes increasingly critical to prioritize cybersecurity measures in order to protect the network and participating organizations from such risks [28].

5.2.2 Addressing privacy and security measures in a blockchain-based environment using cryptographic and non-cryptographic measures

As evidenced by our findings from relevant literature research, privacy and security issues in blockchain-based healthcare applications have been addressed by using a combination of cryptographic and non-cryptographic measures.

5.2.2.1 Smart contracts for access control

Among the various adopted approaches, access control management has been identified as a key strategy. In fact, poor access control measures compromise data integrity, confidentiality, privacy, and availability, thus disrupting healthcare service delivery [210]. Encryption may not be sufficient to protect sensitive data; a ransomware attack may target all data, regardless of encryption. Access control serves as the primary line of defense against data breaches. In this context, what role can the blockchain play in managing system access control?

Smart contracts, self-executing scripts deployed within a blockchain network, are used to manage the access rights of each of the system entities. They are invoked when performing a transaction to execute the term of a contract/procedure on every node in the blockchain [105]. Smart contracts can be designed to provide time-based, role-based, or attribute-based access to data assets, ensuring flexible and secure data access management [83], [88], [93], [104]. For example, Zhao *et al.* implemented a decentralized

fine-grained attribute-based access control through smart contracts to prevent unauthorized access to data resources [104]. They deployed four smart contracts: the user management contract for account creation under which users can store their data resources addresses and access policies; the policy management contract for access policy management by data owners; the access control contract for enforcing the access policies based on user attributes; and the resource management contract for adding and querying data resources and storing the secret share corresponding to users' attributes [104]. In their framework, data owners used a symmetric encryption algorithm to encrypt their personal data before storing the ciphertext in the IPFS system and the returned hash on the blockchain. And to ensure the security of the key, the authors opted for the linear integer secret sharing algorithm, which involves dividing the secret number used to generate the symmetric key into secret shares, with each share linked to an attribute in the access policy [104]. Similarly, Ito *et al.* deployed smart contracts but configured them to give authorized users temporary access to the requested personal health data (PHD). Once the time expired, the permission was automatically terminated [89]. Instead of using IPFS, Ito *et al.* opted for hot and cold storage to store PHD. Cold storage functioned as off-chain and local storage, while hot storage referred to easy-access data storage connected to the Internet like a cloud [89]. As is evidenced in the literature, cryptographic schemes and access control mechanisms are crucial in establishing privacy and confidentiality. However, we will further highlight how selecting a database to store health data holds equal importance.

5.2.2.2 Adopting dual storage with blockchain: a privacy and security management perspective

Off-chain databases, such as cloud repositories, IPFS, graph database, or institutional databases, are used to store health data. Particularly, cloud technology is used as a platform to exchange, store, and monitor information where a centralized virtual database replaces a centralized physical database. But storing sensitive data in the cloud leads to many security and privacy issues [211]. Besides, in a centralized ledger, a single entity controls data. If this entity fails or shuts down abruptly, it will stop processing transactions, and the system will be vulnerable to corruption and fraud. Therefore, IPFS, a decentralized ledger, has been introduced as off-chain storage, an alternative solution to centralized ledgers. IPFS was, for example, used in Chen *et al.*'s blockchain-enabled diabetes disease detection framework to store all patients' health information that was collected via sensor devices [212]. Chen *et al.* also used the blockchain in their framework to store diagnostic results with physiological parameters of a patient after the patient's and

practitioner's approvals [212]. The use of IPFS in storing medical files ensures the scalability of patient data without breaking the decentralization aspect offered by the blockchain. IPFS can also alleviate the storage pressure of blockchain since these processes involve sizeable data. Additionally, it solves the problem of data redundancy by decentralizing the storage of data and using content-addressing to uniquely identify each file. This functionality reduces the overall storage requirements by only storing one copy of each file with the same CID. However, it also brings data privacy issues especially since anyone can use the CID to get data published by other users [213]. Apart from these traditional databases that make it challenging to represent interrelated data structures, we find the graph database valuable in storing linked data, especially after establishing a connection to these data from a blockchain ledger [214]. In fact, the graph database enables knowledge and insights from stored data, but it doesn't address security issues. Given the risks of keeping sensitive data in insecure databases, the question is: what should we store on the blockchain then to ensure secure sharing, handling, and access while maintaining patients' privacy?

To exchange health data while maintaining the integrity and non-repudiation of the datasets, some authors suggested storing the hash pointers or the storage location URL on-chain. Several authors stored the hash of the off-chain record on the blockchain to allow the verification of data authenticity in case a malicious database administrator altered data. Jiang *et al.* added the patient and the hospital signatures alongside the hash of the medical diagnostic report, so both parties were not able to repudiate the administered treatment [69]. Others used the blockchain as a record keeper of keyword ciphertext to help users retrieve the requested data while protecting data security with searchability [68], [69]. Wang, Y., *et al.* proposed a blockchain-based electronic health record (EHR) sharing scheme to ensure data security and privacy preservation among different medical institutions [68]. They utilized cloud-assisted storage to store the EHR ciphertext and brought in consortium blockchain to maintain the EHR indexes. Wang, Y., *et al.* used keyword searchable encryption to ensure data security with searchability and employed conditional proxy re-encryption to enable data sharing with privacy preservation. They included in the data transactions the data provider's signature, in addition to the keyword ciphertext and data owner's account, in order to provide proof of any transaction's validity [68]. A digital signature provides authentication, integrity, and non-repudiation of a message and its source. Storing the data owner/provider signature in a blockchain transaction provides better signature preservation than certification authorities [108]. Once stored on the blockchain, the signature becomes a single shared source of truth where all network actors/nodes can see the same signature, yet none can alter it [108]. Wang, S., *et al.* took advantage of this functionality and

proposed to keep the cloud server signature on-chain, thus entailing patients to verify the correctness of their data sent by the cloud [81]. Their proposal was one way to address the risk of data tampering when storing data in a cloud server.

Another issue involves protecting the privacy of health data stored off-chain in a blockchain-based ecosystem. Xu *et al.* achieved data privacy in their framework by using symmetric encryption algorithms to encrypt off-chain data and implementing fine-grained access control over these data. They stored the hash of the encrypted data on-chain to achieve non-data tampering and placed the symmetric keys on-chain to give the users control over their data, enabling them to add or revoke a doctor at any time [92]. Different from the above work, Zhao *et al.* used the biosensor nodes in the body sensor network (BSN) to generate, back up, and recover the keys used to encrypt data, while the blockchain stored only the ciphertext of physiological data [209]. In this scheme, they solved the data privacy issue by allowing users to control access to their physiological data: users had to ask their biosensor node for the encrypting key and later use this key to restore their data. However, an attacker could compromise the encryption key stored on the biosensor node leading to a potential threat to the patient's data. Encrypting patient data before storing it on the blockchain can help maintain data confidentiality, but this approach might not always be enough to safeguard sensitive information, notably when relying on a third-party authority to set and distribute keys. Secure crypto-key management is crucial for ensuring the confidentiality and privacy of data stored on-chain or off-chain. Mismanagement or loss of private keys can result in unauthorized access to sensitive information [207]. Therefore, in many studies, the cryptographic key was stored on-chain instead of the disadvantageous reliance on a third-party authority to set and distribute [92], [65], [72], [81]. In fact, the encryption keys, once discovered by any attack mechanism, expose the system to data breaches. Using the blockchain to maintain keys will make the management and distribution of keys more secure. Wang, S., *et al.* proposed a health record-sharing scheme in which patients encrypted their health records with a searchable symmetric encryption scheme before storing them in the cloud server [81]. Then, patients saved the symmetric key on the blockchain after encrypting it with an attribute-based encryption scheme. Wang, S., *et al.* enabled privacy by implementing a fine-grained access control that gave patients the right to share their health data with a suitable user through the blockchain [81]. In fact, most frameworks place patients as the sole custodians of their medical records by having them handle access control agreements, as required by the HIPAA privacy rule. In their study, Vanin *et al.* enabled individuals to manage their PHR requests through a data steward as a solution to improve privacy protection and data accessibility in

healthcare institutions [215]. The PHR data was stored on behalf of the individual by the data steward, and any generated data was encrypted using the individual's key. The proposed solution combined IPFS and blockchain to store and distribute PHR data and metadata separately [215]. Additionally, health institutions could temporarily access the requested PHR through a shared data vault after obtaining consent from the individual. To further enhance privacy in data analysis, the solution utilized fully homomorphic encryption techniques, enabling healthcare institutions to perform numerical analysis of encrypted data without compromising individuals' confidentiality [215]. Using homomorphic encryption facilitates data sharing while protecting user privacy [216]. However, it is computationally expensive and has a high latency rate. It is also vulnerable to data inference attacks which can result in encryption key recovery [216].

5.2.2.3 Patient privacy in a de-identified healthcare ecosystem

In addition to implementing access control and data encryption mechanisms, ensuring privacy can be achieved by storing de-identified health data on the blockchain [65], [68], [101]. The blockchain should not disclose data or patients' identities [65]. De-identification, according to the National Institute of Standards and Technology (NIST), involves the removal of personal information from data that could potentially identify an individual [217]. It is a combination of approaches, algorithms, and tools applied to data to reduce costs and privacy risks associated with collecting, sharing, and archiving data [217]. Multiple approaches have been adopted to de-identify health data and safeguard the identity of patients. For example, Uddin *et al.* proposed a framework for continuous patient monitoring that included a patient-centric agent managing a blockchain component to securely store streamed data from body area sensors while preserving privacy [101]. Patient identity was kept private in the system because each person's public key was mapped to a set of symmetric keys by the Security Service Module, which then randomly used one of the linked keys instead of the public key to index the data in the blockchain [101]. Hence, an attacker could not link a patient's record to their relevant physiological data, and the patient's real-life identity remained hidden.

On the other hand, IBM has developed Identity Mixer (Idemix), a cryptographic protocol that allows user authentication without disclosing personal data [218]. Identity Mixer is a zero-knowledge proof (ZKP)

cryptographic protocol that provides anonymity and unlinkability. It has been incorporated into blockchain-based healthcare ecosystems to protect data privacy [219]. It allows participants in the blockchain network to read, update, or share private medical records without revealing their identity. Each participant must use an identity certificate associated with Idemix to perform actions on the distributed ledger [219]. Despite its various benefits, Identity Mixer has limitations, including the inability to issue or use an Idemix credentials with custom attributes and revoke Idemix credentials [220].

Another approach used is anonymization, which involves removing or altering personally identifiable information from datasets, rendering data re-identification impossible [217]. If data can be re-identified, this implies that it has not been effectively anonymized. Data anonymization is adopted to preserve patient privacy and confidential information [221]. While anonymization methods aim to protect against identity, attribute, or membership disclosure attacks, they also have vulnerabilities to certain types of attacks that could expose sensitive information. For instance, the commonly used k-anonymity method protects against identity disclosure but not against attribute disclosure. It is vulnerable to homogeneity and background knowledge attacks [221]. Anonymization methods also face issues such as: privacy risk, data utility, linkage attacks, and data trustworthiness [221]. The selection of an appropriate anonymization method depends on various factors, notably the type of datasets, the nature of the data, and the risk of re-identification. Some models prioritize data utility over privacy, such as: k-map, (1,k)-anonymity, (k,1)-anonymity, and (k,k)-anonymity. Additionally, techniques that involve adding noise to data may affect data truthfulness, leading to a decrease in the overall data utility [221]. Hence, the main challenge in healthcare data anonymization is finding the right balance between privacy and data utility. In a blockchain environment, data anonymization has been introduced to protect patients' identity while handling their data. Chen *et al.* proposed a consortium blockchain-based approach to share medical data securely and protect privacy using k-anonymity, searchable encryption, and attribute-based access control [222]. They chose k-anonymity instead of differential privacy to preserve identity privacy, as the latter can modify the statistical characteristics of the medical data, which may be significant for research or decision-making [222]. On the other hand, k-anonymity can retain the statistical characteristics of data while still providing privacy protection. Moreover, they introduced a system for sharing medical data that employed keyword searchable encryption to maintain data privacy [222]. The encrypted keywords were stored on the blockchain and linked to the encrypted medical data stored on remote clouds. Access control was achieved through a smart contract, which enabled data users, such as research institutions, to either deny access or perform keyword

searches and access the necessary data [222].

Another anonymization technique introduced into the blockchain-based healthcare ecosystem is the ring signature, which guarantees the anonymity of the participants and the information involved in the transaction [204]. The ring signature allows a signer to mix their public key with the public keys of other ring members in a way that makes it impossible to identify the actual signer's key. The identity of the actual signer remains anonymous since the signature can be verified only by using the public keys of the ring members. Plus, the probability of an attacker identifying the signer is only 1/n, where n refers to the number of members in the ring. In their blockchain-based cloud EHR system, Grover *et al.* implemented the anonymous ring signature to preserve patient's data and prevent unauthorized access to these data in the distributed cloud-based EHR system [223]. Deoksang Lee and Minseok Song proposed a privacy-enhanced solution to address the inference problem in blockchain-based health information exchange [224]. Their approach involved using ring signature and stealth address to obscure and conceal sender and receiver addresses. The ring signature employed fake senders to mask a genuine sender, and stealth address used a one-time address as the recipient while providing clues for the genuine receiver to verify the transaction. This approach improves privacy by breaking the link between sender and receiver [224]. However, the solution has some limitations, including slow processing of access requests due to the many elliptic curve operations required for ring signature verification.

Besides anonymization, pseudonymization is another method for ensuring identity protection. The GDPR defines pseudonymization as a method of processing personal data in a way that prevents the data from being directly attributed to a specific individual unless additional information is provided. Therefore, the additional information must be kept separate and subject to technical and organizational measures to ensure that it cannot be used to identify the individual [225]. Some of the basic pseudonymization techniques include encryption and hashing [226]. Cryptographic hash functions can serve not only to secure transactions within a blockchain network but also to derive blockchain addresses [28]. In fact, pseudonymization is commonly used in the blockchain-based healthcare ecosystem through the adoption of blockchain addresses. Blockchain addresses are alphanumeric strings of characters that serve as public identifiers for users in a blockchain network. They are generated by applying the cryptographic hash function to the user's public key, obtained from an asymmetric encryption keygen, along with supplementary information such as version number or checksums [28]. The specific approach for deriving a

blockchain address may vary across different blockchain implementations. For instance, in a permissionless blockchain, users may create numerous addresses by generating multiple asymmetric-key pairs allowing for a certain level of pseudo-anonymity because they are not directly linked to the user's real-world identity [28]. For example, Pham *et al.* used anonymous accounts, notably the Externally Owned Account (EOA) to protect patients' identity in their blockchain-based healthcare ecosystem [94]. In case any attacker gained data access, he would not be able to identify the patients' identity in real life [94]. In fact, EOA accounts are not anonymous but rather pseudonymous because they are linked to a public address on the blockchain. The public address can be used to trace transactions back to the EOA account that has initiated them. De Oliveira *et al.* adopted a similar addressing system for the session key used by the physician to encrypt the patient's medical records. In their blockchain-based schema, the session key was shared on-chain by the patient and addressed with their public key, which implied that the session key's address was derived from the patient's public key, ensuring that the patient's identity remained undisclosed. Privacy and confidentiality were maintained through the patient-centric access control of encrypted medical records and non-disclosure of patient identity in the blockchain [65]. Therefore, De Oliveira *et al.* preserved the privacy of patient data in their framework while ensuring accountability and non-repudiation for physician activities [65]. Likewise, Wang, Y., *et al.* enabled data privacy in their solution by having each entity use its blockchain account, which was anonymous and unlikable to the entity's real identity, to transmit data. The data requester could not get any data owner information, and the cloud server could not deduce the actual identity of the data owner from the EHR ciphertext and re-encryption key [68]. Although the transactions and activities linked to each address on the blockchain can be visible, connecting these actions to the individual behind the address can be difficult without additional information. These studies show that while using blockchain addresses can offer some level of pseudonymity, it is not entirely anonymous because the activity on the blockchain can still be tracked and analyzed to a certain extent.

5.2.3 Key considerations for privacy and security in a blockchain-based environment

While our primary focus in this work is on addressing challenges related to confidentiality and privacy, we must not overlook the importance of maintaining integrity and availability, which are also critical

requirements in any healthcare ecosystem. Hence, multiple factors should be considered in order to maintain privacy and security in a blockchain-based healthcare ecosystem. These considerations include, notably:

- The type of blockchain to be deployed: The choice of the blockchain plays a significant role in preserving privacy [62], [76]. It can be categorized either as permissionless or as a permissioned network:
  - A permissionless blockchain is an open distributed ledger where any node can join the network, and where any two peers can conduct transactions without any authentication from the central agency [105]. A permissionless blockchain does not preserve patients' privacy since any user can see all transactions.
  - A permissioned blockchain includes both private and consortium blockchains. In a private blockchain, a certificate authority controls who can join the network; nevertheless, in a consortium blockchain, a group of preselected members can manage the network. Numerous authors have embraced permissioned blockchains to handle sensitive data due to their capability to offer flexible control over user nodes through access mechanisms [78], [80], [103].
- The off-chain database that will be introduced into the ecosystem: We need to avoid using centralized databases since they are the main reason behind introducing the blockchain as a tool to manage the health information system.
- The cryptographic and non-cryptographic approach to be adopted: It aims at protecting the privacy and security of health data in a blockchain-based ecosystem and includes methods such as data de-identification and access control.

5.3 Proposed blockchain-based healthcare data governance framework: Architecture and design principles

To overcome the limitations of current healthcare systems, we propose an enhanced blockchain-based approach that addresses privacy and security concerns associated with sharing health data among healthcare

organizations. As per the National Commission on Informatics and Liberty (CNIL) and the Data Protection Act 2018, health data are personal information that reveals details about a person's health status, notably their physical or mental health in the past, present, or future, and the healthcare services they have received [227]. They also include the information collected when registering a person for health services, such as a patient identifier. Even though the concept of health data has broadened in recent years to include a wide variety of data, we focus on those that are health data by nature. Therefore, we have worked on the health record. A health record is a vital tool that healthcare professionals use to record, track, and share information with other care providers regarding the service that has been or is to be provided to a patient [228]. It consists of two sorts of data: clinical and administrative. Clinical data include information related to patients' health, notably their medical history, diagnosis and treatment plans, report of consultation or hospitalization. Administrative data include patients' demographic and financial information [228]. Because confidentiality and privacy are our top priorities, we must ensure that health information is not exposed to unauthorized parties and that patient-identifiable information is protected. In addition, dealing with health data requires compliance with regulatory guidelines that dictate how health information should be used and disclosed. We distinguish notably:

- HIPAA: a federal US law that mandates the development of national standards to protect sensitive patient health information from being disclosed without the patient's consent or knowledge. The privacy rule aims to protect individual's health information while allowing the flow of health information needed to provide high-quality healthcare and well-being [229].
- GDPR: a regulation in EU law on data protection and privacy. It aims to enhance individuals' control and rights over their data. Two of these rights are the Right to Rectification and the Right to Erasure [230].

Data governance is a set of practices and policies designed to manage the availability, usability, and security of data during its life cycle. Even though data governance includes many key aspects, our proposal addresses the security and privacy aspects of patients' data when shared among hospitals and between hospitals and pharmaceutical companies. We will focus on three healthcare activities: patient care, emergency care, and clinical research. Since patients must initially register to benefit from healthcare services, we have included in our study the administrative activity involving 'healthcare registration'. Our proposal aims

to empower patients by giving them more control over their health data while ensuring that healthcare companies respect their privacy and preferences when handling or sharing their data. Hence, to achieve our aim, we have implemented data segregation, data de-identification, and robust security measures, such as access controls and encryption, to protect data from unauthorized access. However, a uniform handling of health data can lead to inadequate protection of sensitive information, thereby increasing the risk of unauthorized data access. For this reason, prior to implementing any security measures, data classification becomes crucial as it enables a more targeted approach to safeguarding sensitive information [231].

5.3.1 Data classification

According to the International Standards Organization (ISO) and the NIST, data classification schemes allow efficient information management and security based on their criticality, value, and potential risks associated with disclosure or damage [231]. In general, organizations classify data following their specific needs and objectives. Healthcare information is classified as confidential because unauthorized disclosure can seriously affect the healthcare organization and the patient [232]. Any activity linked to health information must be controlled, audited, and monitored [232]. In our study, we have decided to adopt a data classification approach based on the data's relevance to their intended use, assisting healthcare organizations in wisely allocating resources for cost-effective data safe-guarding. Consequently, evaluating the significance and relevance of data for administrative purposes or healthcare activities, such as patient care, research, and public health, enables tailored security measures and access control.

On the other hand, we are aware that in the context of large-scale medical research studies, protected health information (PHI) can be shared without compromising patient privacy and without requiring individual authorization. This is feasible when the information undergoes effective de-identification, ensuring there is no risk of re-identification [203]. HIPAA has validated two de-identification methods: Safe Harbor and Expert Determination. The Safe Harbor method involves removing 18 specified identifiers of PHI, including both direct and indirect identifiers, to reduce the risk of re-identification of patients [203]. Unlike Safe Harbor, Expert Determination relies on the knowledge and expertise of a qualified statistical expert to determine the risk of re-identification and ensure that it is not likely for someone to identify an individual even when this information is combined with other available data [203]. As part of our approach,

we have implemented de-identification by separating patients' identifiers from the data and eliminating any direct connections or links between them. Therefore, when classifying data, we consider both: the criticality of health information, especially in emergencies, and the 18 identifiers that should be removed, as per the Safe Harbor de- identification method. Having identified the distinct healthcare record components, we have proceeded with data classification using these proposed categories:

1. Critical: describing data that are readily available and essential for addressing urgent medical needs.
2. Mandatory: describing data that are essential for administrative purposes and various healthcare activities, such as patient care, clinical research, emergency care, etc.
3. Optional: describing data that are advantageous but not essential for administrative purposes and various healthcare activities.
4. Restricted: describing data that should not be disclosed to individuals carrying out the healthcare activity or administrators due to privacy concerns or legal requirements.

Table 5.1 shows the data classification we have applied to the healthcare record elements, based on their relevance in achieving a healthcare activity, notably patient care, emergency care, clinical research, or a healthcare administrative activity. Patient care encompasses a wide range of healthcare services, from initial consultations to routine immunizations and checkups to higher levels of specialty care such as hospitalization. For each data element, we have additionally checked whether it falls under one of the 18 identifiers specified in the Safe Harbor de-identification method. Based on the collected information, we have ascertained the most fitting accessibility, storage, and cryptographic techniques for each category, ensuring the desired level of security. According to the Safe Harbor de-identification method, the identifiable data that must be removed from health data are names, addresses, phone numbers, email addresses, IP addresses, social security numbers, medical record numbers, health plan beneficiary numbers, device identifiers, certificate/license numbers, account numbers, vehicle identifiers/serial numbers, website URLs, full-face photos, biometric identifiers, all element of dates (birthday, admission or discharge dates, etc.), and any unique identifying numbers, characteristics, or codes [203]. It is important to note that most of these items are not typically included in our table, especially since we have excluded the Internet of Medical Things (IoMT) from our study and because some of the identifiable are irrelevant in a healthcare

record. Besides, some healthcare record data elements are not among the 18 identifiers reported by the Safe Harbor method but can still identify a patient unless omitting the identifying information. For instance, we cannot ascertain a patient identity from a laboratory test result if it lacks the patient's name, birth date, or healthcare record ID. These data elements are classified in our table as potentially identifying information but only in certain conditions. In addition, some healthcare record data elements, such as gender, blood group, occupation, etc., cannot identify a person unless correlated with additional healthcare record data. These data elements are classified in our table as non-identifying information but only in certain conditions. It is also important to note that even though healthcare data are mandatory to achieve a successful clinical research activity, we have chosen to classify them as optional throughout the categorization process. In fact, the type of data collected during a clinical research activity is driven by the research objectives and the patient's willingness to participate in such studies.

Table 5.1: Data classification of the healthcare record elements based on achieving the following healthcare activities: patient care, emergency care, clinical research, and healthcare registration.

Healthcare Record	**Potentially identifying information (Safe Harbor)**	**Patient care**	**Emergency Care**	**Clinical Research**	**Healthcare Registration**
1. Patient Demographics					
Full-face photo	Yes	Optional	Optional	Restricted	*Mandatory*
Name	Yes	Optional	*Mandatory*	Restricted	*Mandatory*
Address	Yes	Restricted	Restricted	Restricted	*Mandatory*
Healthcare record number	Yes	*Mandatory*	*Mandatory*	Restricted	Restricted
Telephone number	Yes	Optional	Optional	Restricted	*Mandatory*
Email address	Yes	Optional	Optional	Restricted	*Mandatory*
Date of Birth	Yes	Optional	*Mandatory*	Restricted	*Mandatory*
Gender	No[1]	*Mandatory*	Critical	Restricted	*Mandatory*
Marital Status	No[1]	Optional	Optional	Restricted	*Mandatory*

Table 5.1 – *Continued on next page*

[1]The data element does not identify a person if not correlated with additional information from the healthcare record.

Table 5.1 – *Continued from previous page*

Blood Group	No[1]	Optional	Critical	Optional	Restricted
Emergency Contact Information	Yes	Optional	Critical	Restricted	*Mandatory*
Occupation	No[1]	Optional	Optional	Restricted	*Mandatory*
Company name, address, and phone number	No[1]	Optional	Optional	Restricted	*Mandatory*
2. Financial Information					
Social security number	Yes	Restricted	Restricted	Restricted	*Mandatory*
Health plan beneficiary number	Yes	Restricted	*Mandatory*	Restricted	*Mandatory*
Insurance provider's name, address, and phone number	No	Restricted	*Mandatory*	Restricted	*Mandatory*
Insurance policy number, group number, effective and expiration dates, and any applicable policy holders or beneficiaries.	Yes	Restricted	Restricted	Restricted	*Mandatory*
Claim and billing information	Yes	Restricted	Restricted	Restricted	*Mandatory*
3. Medical History					

Table 5.1 – *Continued on next page*

Table 5.1 – *Continued from previous page*

Medical Conditions	Yes[2]	*Mandatory*	Critical	Optional	Restricted
Surgical History	Yes[2]	*Mandatory*	*Mandatory*	Optional	Restricted
Family Medical History	Yes[2]	*Mandatory*	*Mandatory*	Optional	Restricted
4. Medication History					
Current Medications	Yes[2]	*Mandatory*	Critical	Optional	Restricted
Prescription History (Start date, End Date, Symptoms, etc.)	Yes	Optional	Optional	Optional	Restricted
Allergies to Medications and Adverse Reactions	Yes[2]	*Mandatory*	Critical	Optional	Restricted
5. Treatment History					
Date of Treatment or Date of hospitals admission and discharge	Yes	Optional	Optional	Restricted	Restricted
Treatment and Interventions details	Yes[2]	Optional	Optional	Optional	Restricted
Treatment Outcomes	Yes[2]	Optional	Optional	Optional	Restricted
6. Laboratory and Diagnosis Tests					

Table 5.1 – *Continued on next page*

[2]The data element is not among the 18 identifiers reported by the Safe Harbor method. But it can include identifying information such as the patient's name, date information, healthcare record number, or any given unique identifiers by a healthcare facility.

Table 5.1 – *Continued from previous page*

Blood Tests, Imaging Studies (MRI, X-ray, etc.), Electrocardiogram, etc.	Yes[2]	*Mandatory*	Critical	Optional	Restricted
Pathology Reports	Yes[2]	*Mandatory*	Critical	Optional	Restricted

After classifying the health record data, the next step involves selecting the appropriate blockchain and off-chain storage to integrate into our ecosystem.

5.3.2 Blockchain selection

We have introduced blockchain technology into our ecosystem to improve information sharing among healthcare stakeholders while maintaining patients' privacy and granting them control over their data. The blockchain empowers patients by allowing them to choose what information to share and with whom, thus ensuring their rights are respected. A permissioned blockchain, with a trusted identity provider managing the identity of all network users, is the most appropriate in our case. The identity provider is trusted to maintain access control within the network and the users' rights to participate in the consensus or validate a new block [28]. We have adopted the Hyperledger Fabric, an open-source permissioned blockchain that restricts network access to authorized participants. Hyperledger Fabric is one of the most mature technology projects in the Hyperledger platform and the most popular due to its scalability and rich documentation. Fabric is highly scalable due to its flexible and modular design. It allows pluggable consensus protocols, including BFT, and uses CouchDB as a state database to store data. CouchDB provides scalability, enabling Hyperledger Fabric to handle expanding workloads and accommodate large-scale deployment. In addition, Fabric offers developers multiple programming languages to write chaincodes, such as JavaScript, TypeScript, Golang, Java, and Node.js [30]. A chaincode is a piece of code that defines the business logic and rules governing the interactions and transactions in the Hyperledger Fabric network [30], [179].

Moreover, Fabric includes a membership service provider (MSP) that manages users' IDs and authenticates all network participants. A Hyperledger Fabric blockchain network can be governed by multiple MSPs, thus providing [30], [179]:

- Modularity in membership operations because different MSPs have autonomy and independence in managing their membership functions.
- Interoperability across diverse membership standards and architecture since it allows participants with different membership architectures to collaborate and transact with the Hyperledger Fabric. The default MSP implementation in Fabric utilizes X.509 certificates as identities and the Public Key Infrastructure (PKI) model.

Each participant has a unique digital certificate which includes their public key and identifying information. By accurately identifying and attributing actions to specific network participants, Hyperledger Fabric promotes accountability and improves system integrity. In addition, Hyperledger ensures confidentiality by leveraging its channel architecture and private data feature [30], [179]. Channels in Hyperledger Fabric create sub-networks within the overall network, where specific participants can view a defined set of transactions. Hyperledger Fabric achieves sharding through channels, thus improving scalability. In fact, each channel acts as a shard and has its ordering service responsible for ordering transactions and enabling a secure transaction processing [233]. If an ordering service can't handle all the transaction load of its current channels, a new ordering service can be spun up to accommodate additional channels [233]. In addition, every channel possesses its respective ledger and a distinct collection of chaincodes [30], [179]. A channel includes multiple organizations, each having its role in the network and managing its different nodes, notably the peer nodes. A peer node maintains a copy of the ledger and executes transactions based on the defined chaincodes. Furthermore, channels are configured with access policies that control access to the channels' resources, notably chaincodes, transactions, and ledger state. Hence, the privacy and confidentiality of information are preserved within the channel nodes [30], [179].

At another level, private data collection allows for the confidential storage and sharing of data among a subset of participants within a channel. Figure 5.1 shows a Peer's ledger with enabled private data collection consisting of two elements: the actual private data stored in a private database on the peer known

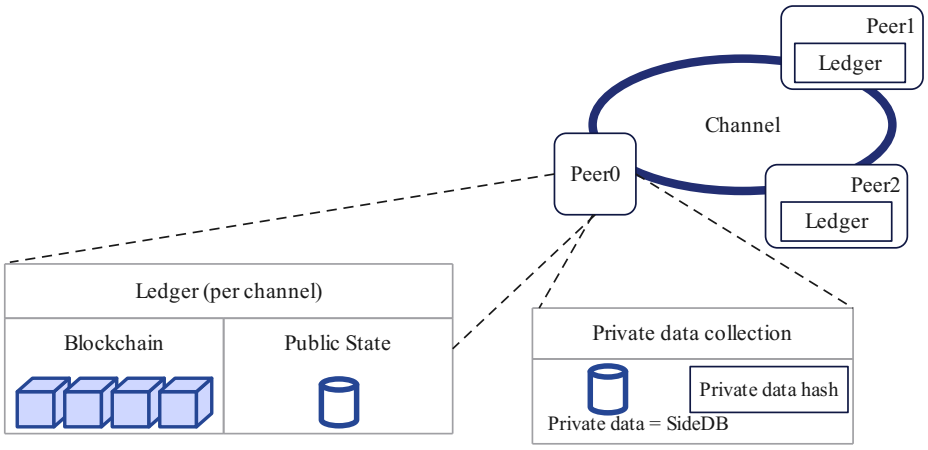


Figure 5.1: Peer's ledger with enabled private data collection

as "SideDB" and a hash of that data stored publicly on the channel [30], [179]. The data hash has two objectives: it validates the current state before altering private data in the SideDB of authorized peers; it serves as an auditable proof of the transaction [30], [179]. Moreover, the transaction flow involving private data is modified to protect the confidentiality of transmitted data. In fact, the private data collection feature keeps private data confidential from the ordering service since all private data are sent peer-to-peer via gossip protocol to the subset of peers authorized to see it. The ordering service and anyone querying the blockchain ledger only sees the data hash [30], [179]. In addition, when a chaincode referencing private data collections is invoked, the private data are passed via a transient field in the transaction proposal. The transient field allows data to be included in a transaction without being permanently stored in the ledger. Data are treated as temporary and are only made available to the endorsing peers during the transaction endorsement process. Endorsing peers simulate the transaction and store the private data temporarily in a transient data store local to each peer [30]. As per the collection policy, private data get distributed to authorized peers using the gossip protocol. These authorized peers validate then the private data against the hashes in the public block and proceed to commit the transaction [30]. Consequently, the private data are transferred to their private state database before being removed from the transient data store. As a result, to accomplish data privacy goals and in line with our ongoing proposal, we should take the following steps:

- Partition the healthcare network into channels, with each channel representing a subset of the healthcare actors that are authorized to see the data handled by the chaincodes deployed to that

channel.

- Use private data collection, which allows patients to share specific information with the physician for a certain period while recording the data hash on the public ledger as evidence of the transaction occurrence.

- Control data access based on the attributes delivered to the network participants by building access control into the chaincode logic.

- Implement data encryption at different levels in accordance with privacy and security requirements of the data and the regulatory framework governing the data (application, database, file system, or network level).

5.3.3 Off-chain storage selection

As part of our strategy, we have adopted a dual storage architecture to efficiently handle extensive data volumes without compromising the performance of the blockchain. We have chosen the InterPlanetary File System (IPFS), a decentralized file-sharing protocol, as off-chain storage [234].

IPFS is a content-addressed peer-to-peer distributed file system that seeks to connect all computing devices with the same system of files [235]. In IPFS, a node or peer refers to a single instance of IPFS software running on a machine. One machine can run one or more IPFS nodes if required. Deploying several instances of IPFS results in a more robust and decentralized network, which enhances network performance and resilience. Each IPFS node in the network has a unique identifier, also known as Peer ID, which is the cryptographic hash of the node's public key [234], [235]. By default, when adding content to one IPFS node, the content is stored as blocks in that node's local repository and can be retrieved from that node using the content identifier (CID). IPFS uses directed acyclic graphs (DAGs) to store data in a way that is both efficient and secure [234], [235]. Data are broken down into blocks, and each unique block has a unique CID, which is the content address of its data and is generated based on its content's cryptographic hash. The blocks are interconnected in a graph structure, with each block pointing to its parents and children, forming a tamper-resistant data chain [234], [235]. The content is not visible to other

IPFS nodes on the network until it is announced and shared with them. IPFS uses Kademlia, a distributed hash table (DHT), to find the closest nodes in the peer-to-peer network storing the desired data [234]. The DHT enables efficient content discovery and retrieval. It uses Libp2p, a library peer-to-peer networking framework, to establish connectivity between IPFS nodes [234]. It stores records in a key-value pairs format that indicates which peers have which blocks. When users need to locate a particular file or content within the DHT, they can utilize the Content Identifier (CID) as the key for retrieval [234]. The DHT then returns the network locations (PeerID) of nodes that contain that content, allowing the user to retrieve the desired file from those locations. Unlike centralized systems, Kademlia can be visualized as a single large table split across all peers. Each IPFS node maintains a local copy of the DHT and participates in the DHT by storing a subset of the overall data. Furthermore, when a specific content gains popularity within the network, its discoverability improves as numerous peers are able to provide it [234]. Figure 5.2 represents the IPFS architecture and outlines the steps of adding a file larger than 256 kilobytes. As previously stated, the file is divided into smaller chunks (< 256). Each chunk is hashed to generate a unique CID. These CIDs are arranged to create the Merkle DAG. The resulting base CID is added to IPFS Node A. When a node joins the IPFS network, it connects to other nodes and begins requesting and serving blocks to other nodes using BitSwap, the block-exchange protocol [234], [235]. A block sent via a BitSwap is prefixed with the CID version, multicodec, and multihash. Once a block is received, these parameters are used to hash the data and calculate the CID. If the calculated CID matches a value in the node wantlist, it is saved to the node's local repository. Otherwise, it is dropped [234]. CIDs must be pinned to an IPFS node on the network to guarantee that they are constantly accessible or available, ensuring that the content is stored locally on the machine [234]. Content that has not been pinned or referenced is deleted during garbage collection. The garbage collection runs periodically to free up disk space and ensure the IPFS repository is not cluttered with unused data [234]. By default, it runs every hour. However, it is possible to set the time interval according to the storage requirements [234].

Since our use case scenario involves using IPFS to store patient healthcare data, which will grow over time, it is crucial to set up an IPFS cluster within our healthcare ecosystem. IPFS cluster can improve scalability by distributing the load of storing and serving data across multiple nodes [234]. It allows faster data access for users since the demand is not concentrated on a single node, thus prevents performance bottlenecks [234]. An IPFS cluster also ensures data availability and redundancy within the internal network and facilitates the management of multiple IPFS nodes. Therefore, files are added to one node and

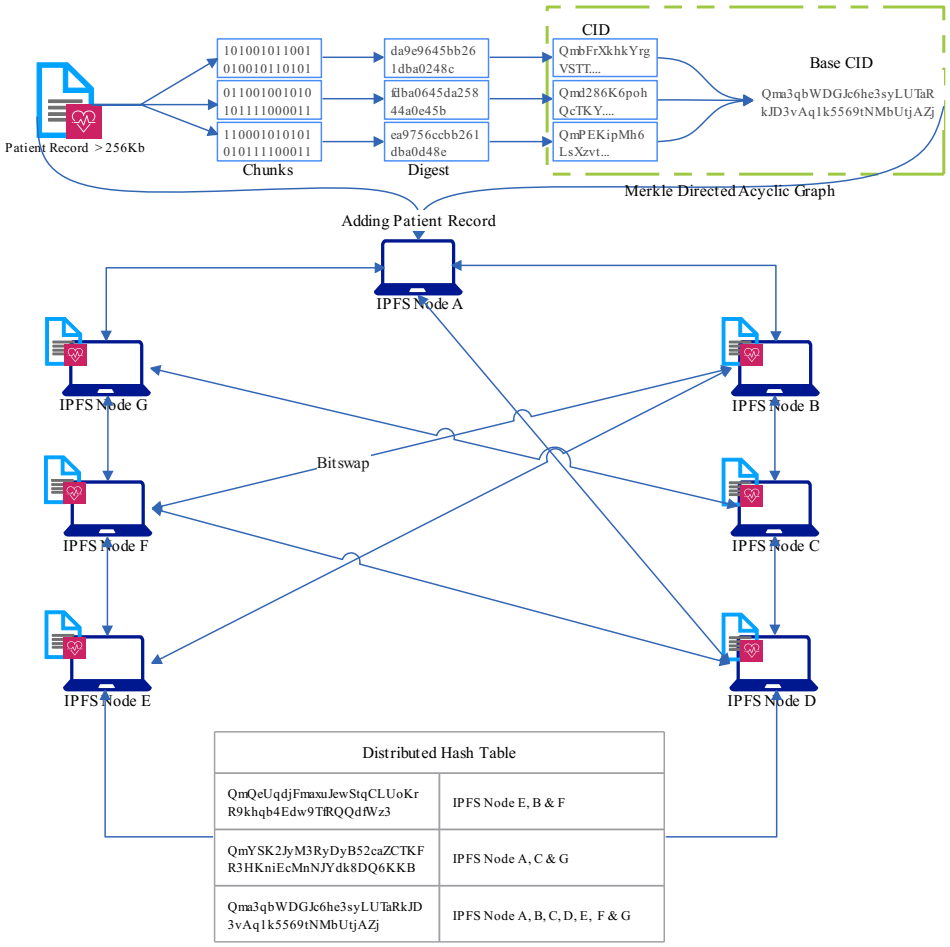


Figure 5.2: High level view of IPFS Architecture

replicated across other nodes in the cluster according to the chosen replication factor. A high replication factor enables reliability, increases resiliency, and improves network performance [234].

From a security perspective, the cryptographic hash in IPFS provides a security layer by verifying the authenticity and integrity of files, making it difficult for malicious actors to tamper with or delete them [234]. Nevertheless, even though IPFS is considered immutable since any modification in the content results in a different identifier (CID), it is still possible to delete content from the IPFS at some point. In fact, deleting a block associated with a CID from a node is possible if the CID has not been replicated to other nodes. The deletion is a local operation like the pinning process [234]. In case it has been replicated, we need to contact other nodes that could have a copy of the content and request that they delete it. This feature in IPFS helps fulfill patients' rights under the GDPR. On the other hand, it is essential to encrypt all

the content added to IPFS even when deploying a private IPFS, because IPFS does not inherently provides data protection in terms of privacy and confidentiality. Encrypted content, even when possible to track, cannot be read by network users without the decryption key. Therefore, it is necessary to ensure that the decryption keys are highly guarded. Additionally, we must implement read-write access control on the IPFS to ensure that only authorized users can access the data.

5.4 Putting the proposed data governance framework into action: A tailored network configuration for ensuring privacy and security in a healthcare ecosystem

We have identified in our healthcare ecosystem four actors and their respective roles:

- The patient, responsible for seeking treatment and controlling access to their healthcare record by sharing it with the proper stakeholders.
- The physician, responsible for diagnosing, treating, and monitoring patients.
- The clinical researcher, responsible for designing, implementing, and monitoring clinical trials to determine the efficacy of medical treatment.
- The administrator, responsible for creating network participants' demographics, particularly those of patients, physicians, and clinical researchers. Each organization featured in our network (Figure 5.3) has its administrator. We cite, notably, the healthcare authority administrator, the hospital registrar, and the pharmaceutical company administrator.

A high-level architecture of our proposed system is depicted in Figure 5.3. Four organizations are collaborating, namely the healthcare authority, two hospital organizations, and the Research & Development department of a pharmaceutical company. Each of these organizations has its membership service provider (MSP) with its certificate authority (CA). We have chosen to integrate the healthcare authority as a third

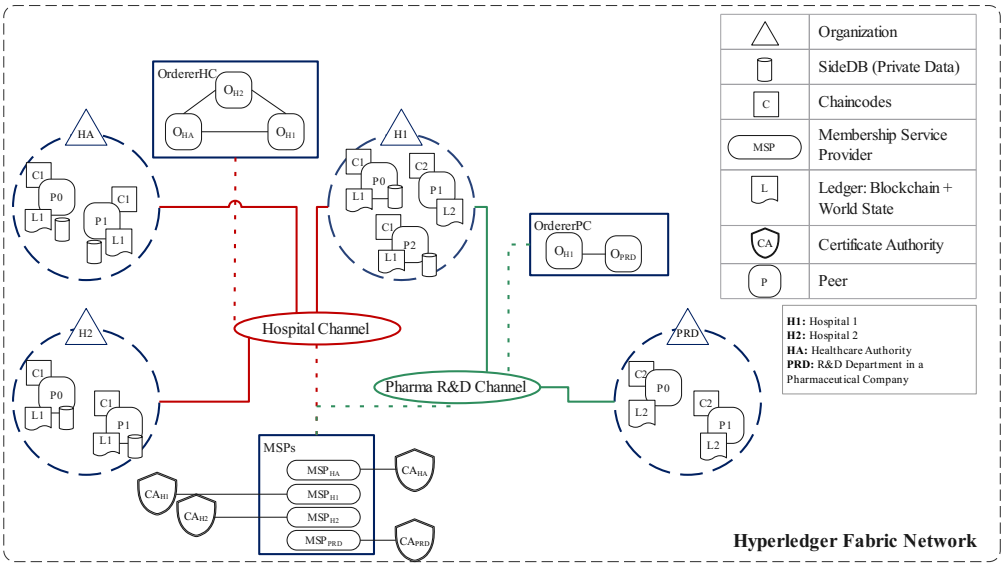


Figure 5.3: Hyperledger fabric-based network topology for managing patients' data between hospitals and between hospitals and pharmaceutical companies' Research & Development department

organization into our network because the goal of segregating the management of data requires that no single actor has complete control and unrestricted access to all data. The healthcare authority is an organization that can be represented by a country's Ministry of Health or any global health organization, and that is in charge of managing patient demographics and financial information. Although we have focused on a specific use case involving hospitals and pharmaceutical companies, our solution holds potential for broader applications. It can be expanded to include a wide range of stakeholders within the healthcare domain (such as health insurance companies, pharmacies, etc.) while being adaptable to suit other domains beyond healthcare.

Given the sensitive nature of health data and the necessity to handle these data with utmost confidentiality, it is essential to segregate the different interactions involving the system actors. In a healthcare environment where patient care, emergency care, and clinical research intersect, segregation of data management is essential. The data handling requirements and the conditions that must be maintained vary significantly across healthcare activities. This is due to the varying levels of sensitivity associated with different types of healthcare record data and the different levels of access required by the various healthcare activities. Patient care data, for example, must be handled with utmost confidentiality whereas emergency care data may need to be shared more quickly and easily. On the other hand, clinical research data must

be thoroughly de-identified before being shared with stakeholders such as researchers and pharmaceutical companies to remove all personal identifiers.

Therefore, we have implemented two channels within the Hyperledger Fabric framework, as shown in Figure 5.3:

- The hospital channel, hosting all the exchanges among the patients, physician, emergency physician, healthcare authority administrator, and hospital registrar to enable better care coordination.
- The pharma R&D channel, hosting all the exchanges among the patient, physician, and clinical researcher to improve the quality of care through the development of new treatments for diseases. It is important to note that a clinical trial is led within a hospital by physicians known as principal investigators in collaboration with the R&D department of a pharmaceutical company. However, to narrow the scope of our research, we have merged the roles of medical practitioner and principal investigator under one actor: the physician.

By creating two channels, we have segregated the healthcare activities to enhance data handling and meet the requirement of each environment. Each channel has its own ledger, access controls, policies, and tailored workflow to meet the needs of each of its actors' activities. Moreover, we have used the data classification shown in Table 5.1 to assess how data will be handled in our blockchain-based ecosystem. By considering the data type and its relevance to a specific activity, we can make informed decisions about the storage and the cryptographic and non-cryptographic techniques to adopt.

5.4.1 Hospital network

The hospital channel represents the first network in our ecosystem that connects multiple hospitals together for the purpose of sharing patient data. The different interactions in this network are mainly related to the patient visit cycle depicted in Figure 5.4.

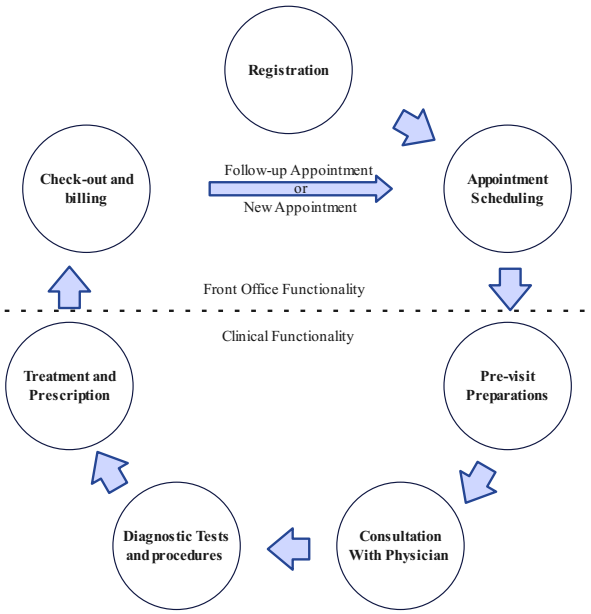


Figure 5.4: Patient Visit Cycle

We address the interaction between a patient and a physician in two contexts: patient care and emergency care. They both fall within the medical services scope, which includes diagnosis, treatment, and management of individuals' health conditions. The only difference is that patient care is preventive and involves long-term health care while emergency care focuses on the immediate treatment of urgent medical issues.

Therefore, we segregated patient health record data across three distinct storage locations, as shown in Table 5.2:

- IPFS, where extensive health record data are stored. Two distinct private IPFSs are deployed in our ecosystem. One IPFS stores patients' administrative data, and the other stores de-identified patients' clinical data following the HIPAA privacy rule. We have used distinct symmetric keys to encrypt the administrative and clinical data stored on IPFS. We will cover the specifics of these keys later in the text.
- Private state database, where small sets of metadata related to patients' clinical data (the part stored on IPFS) are stored, considering that private state database is not designed to handle large datasets [233]. Even though it is possible to enforce read or write access to private data by setting the

collection fields *memberOnlyWrite* and *memberOnlyRead* to true, we have opted to set them to false to implement a fine-grained access control mechanism within the chaincode logic. By setting these variables to false, the read and write access are no longer limited to the collection policy members.

- Ledger, where critical health record data needed to provide urgent care are stored. These data cannot identify a patient if not linked to additional information.

Before proceeding with the network interactions' details, it is essential to understand how the data classification in Table 5.1 affects the access control configuration. Access control is implemented at the chaincode level, which gives us the ability to manage client access to public or private state data. For example, we can decide who can query or write within a private data collection based on the client attributes embedded in the client certificate. Table 5.2 provides an overview of the different private data collections within our ecosystem, with an insight into how data are distributed across multiple storage locations and the permission rights for each network actor. In brief:

- Critical data are stored on-chain or in a private data collection without encryption to help emergency physicians retrieve the essential data elements required to provide immediate intervention.
- Mandatory data are stored in private data collection or IPFS. These data are provided by a patient to an attending physician before a medical appointment, or by an emergency contact to an emergency physician when the patient is unconscious.
- Optional data are stored on private data collection or IPFS. These data are shared by a patient with a physician upon request, which may introduce a certain level of delay.
- Restricted data are stored on private data collection or IPFS. Sharing these data is not allowed.

5.4.1.1 Patient care

In a patient care context, we will present our solution based on the sequential steps highlighted in Figure 5.4. These steps represent the journey every patient undergoes when seeking medical care.

5.4.1.1.1 Registration

In order to access healthcare services through our framework, patients should complete the administrative registration activity. As previously highlighted, we have segregated the administrative data from the clinical data to ensure the privacy of patient data, each set handled by a different network actor. Therefore, each patient has two distinct identifiers: an administrative ID assigned by the healthcare authority administrator and a healthcare record ID (aka healthcare record number) assigned by the hospital registrar. The administrator of the healthcare authority collects patients' demographic and financial information - except for their blood group and healthcare record ID – and stores these data in the private IPFS managed and maintained by the healthcare authority. A private data collection named **PatientDemFin** is created at the healthcare authority level. This collection stores patients' administrative IDs issued upon registration to our framework, associated with the IPFS URL representing the address of patients' administrative data stored on IPFS. We have chosen not to include the patient blood group since this information is irrelevant to healthcare administrative activities. And by concealing the healthcare record ID from the healthcare authority administrators, we prevent them from linking multiple IDs to a single patient.

Because patients may seek medical care from multiple hospitals, keeping track of their visits and interventions is essential to provide high-quality care. As a result, we must adopt a centralized metadata repository, which enables efficient data retrieval, regardless of its physical location, and helps patients maintain control over all their data. Therefore, during the registration, patients must choose the hospital that handles their healthcare record metadata, also known as the primary hospital. Each hospital has a private data collection dedicated to its patients, named **Patient-hospID**, with hospID representing the hospital ID (e.g., Patient-hosp1). This approach enhances security since it limits patient access to the private data collection that detains their healthcare record metadata.

In order to interact in the Hyperledger Fabric network, each participant must have a certificate and a private key. As shown in Figure 5.3, each organization has its CA responsible for managing network actor certificates and registering identities. During the registration process, we suggest two different approaches, with the main difference residing in the organization in charge of issuing the patient's certificate. The certificate could be issued based on the patient's administrative ID or healthcare record ID. Since we are focusing on a narrowed use case scenario, the rightness of these approaches depends on the different healthcare actors interacting and the security constraints of the addressed case study.

- Approach 1

  The healthcare authority CA generates a certificate for each patient based on their administrative ID and not their name. The certificate includes two distinctive attributes: the primary hospital ID chosen by the patient and the user role, 'patient'. These attributes are essential for access control management. Each patient has the enrollment keys required to transact within the Fabric-based healthcare ecosystem. Simultaneously, the healthcare authority administrator shares with the primary hospital registrar the administrative ID of the newly enrolled patient. Regarding the patient information stored on the healthcare authority IPFS and as shown in Table 2.1, four data components are mandatory in emergency cases: patient's name, date of birth, health plan beneficiary number, and insurance provider contact details. It is necessary to encrypt all administrative data to maintain their confidentiality. Each patient has three symmetric keys to encrypt the administrative emergency data, the patient's demographic data, and their financial information, respectively. The healthcare authority administrator encrypts these keys by the patient public key before storing them in the **PatientDemFin** collection. One disadvantage of this approach is that the healthcare authority can track the patient on the network.

- Approach 2

  The certificate authority of each hospital generates the certificate for its patients based on their healthcare record ID and not their name. The certificate includes two distinctive attributes: the hospital ID and the user role, 'patient'. In this approach, the healthcare authority encrypts patients' demographic and financial information with a symmetric key before adding these data sets to its private IPFS. They are encrypted with a symmetric key owned by the healthcare authority. One advantage of this approach is that the healthcare authority cannot track patients over the network because they are unidentifiable. The patient should not interact with the healthcare authority on the hospital network to avoid any possible identification. Therefore, any patient's request to update their administrative data is carried outside the hospital network. One disadvantage of this approach is that the healthcare authority has complete control over patients' administrative data. In the case of an emergency, the healthcare authority administrator needs to share the mandatory data with the requester while being compliant to the network policy. Hence, the challenges introduced in this approach pertain to how patients update or monitor their shared data.

After enrolling patients, the next step is enrolling physicians in the Fabric-based healthcare ecosystem. In our proposition, a physician is affiliated with one hospital. As a result, each hospital registers its physicians, and a certificate is generated by the hospital CA for each physician based on their physician ID rather than their name. The certificate includes four distinct attributes: medical specialty, medical subspecialty, hospital ID, and user role - that is, physician or emergency physician. In each hospital, a private data collection named **Physician-hospID** is created, with hospID representing the hospital ID. This collection stores physicians' information, notably full name, physician ID, medical specialty, contact information, and their availability. While each physician and hospital registrar have the right to access the **Physician-hospID** collection of the hospital they belong to, only the physician can write into this collection.

5.4.1.1.2 Appointment scheduling

The hospital registrar must perform the following task for each new patient, regardless of the approach adopted at the registration level :

- Assign the patient a healthcare record ID: Upon receiving the administrative ID of the new patient, the hospital registrar provides them with a healthcare record ID before adding it alongside the administrative ID in the **Patient-hospID** PDC.
- Notify the hospital's emergency physicians of the new patient's arrival by adding the patient's healthcare record ID to the **Emergency** PDC. An emergency physician then assigns the patient two emergency IDs. These IDs are added to the Emergency PDC along with the gender, blood group, and emergency contact information of the patient. The first emergency ID (emergencyID1) is designed to be easily handled by the patient, enabling better patient management by the emergency medical team. This ID can be presented as a physical card, a wearable device like a bracelet, or a digital ID kept on the patient's smartphone. It enables the emergency physicians to access the second emergency ID (emergencyID2), which in turn allows them to access the critical data stored on the ledger needed for providing first aid to the patient. In that way, we maintain patient privacy, preventing any network intruder from identifying a patient from their emergency ID.

Once a patient is registered, the hospital registrar's responsibility is to book an appointment and connect

the patient with the preferred physician by providing the patient with the physician ID, appointment date, and time. This process can be executed off-chain or on-chain. In the on- chain scenario, the hospital registrar invokes the *putNewAppointmentPrivate* function, which takes the collection name, patient's healthcare record ID, physician ID, and appointment date/time as arguments. This function stores the appointment details of hospitals' physicians in the hospitals' **Appointment-hospID** PDC, with hospID representing the hospital ID. Physicians are granted read-only access to the **Appointment-hospID** PDC of the hospital they are associated with, while patients have read-only access to the **Appointment-hospID** of all hospitals. In our solution, we have adopted the off-chain approach to keep it brief.

5.4.1.1.3 Pre-visit preparation

Pre-visit preparations are the steps a patient takes before their appointment, which can help improve the efficiency and effectiveness of the visit and ensure that the patient receives the best care possible. They involve gathering any previous healthcare records or bringing the result of a diagnosis or lab test requested by the physician in a prior consultation. By sharing, in advance, all relevant medical records, the patient gives the physician valuable insight into their medical history, which helps the latter gain time in assessing the situation and providing a more tailored treatment plan. The patient adds to the **Patient-Physician** PDC the health record data considered mandatory or optional for medical consultation, along with the physician ID and patient healthcare record ID. The **Patient-Physician** PDC is shared among all network hospitals. Hence, all added data is encrypted with the physician's public key except for the patient and physician identifiers.

The data stored in the **Patient-Physician** PDC is purged after a specified number of blocks, represented by blockToLive, a data collection parameter. The purpose of purging data is to maintain the privacy and security of sensitive information by not retaining it indefinitely. The blockToLive value chosen should allow the physician enough time to evaluate the shared data before it becomes unavailable on the network. As for the **Patient-hospID** and **PatientDemFin** data collections, the blockToLive is set to 0 because the patient's metadata must be stored indefinitely unless the patient decides otherwise.

5.4.1.1.4 Consultation, Diagnostic, Treatment, and Prescription

After treating the patients, physicians are responsible for accurately documenting the consultation details

in the patients' healthcare records. It includes updating or adding relevant examination findings, notably diagnostic test results, treatment plans, treatment progress, and drug prescriptions. Since each health data element is handled differently, a physician must store collected data in their respective locations, following our network policy. For example, adding a treatment consists of adding the following elements: hospital ID, date of treatment, purpose, outcomes, and treatment plan, which includes details on the procedures, therapies, or medication involved. In such a case, the physicians will add the treatment plan, purpose, or treatment outcomes without identifying details (date, hospital ID (location), patient identifier, etc.) to the IPFS shared among all the hospitals in the network. These data are encrypted with a symmetric key, previously encrypted with the physician's public key and shared on the **Patient-Physician** PDC by the patient. Afterward, the physicians write the private treatment details in the private data collection **Patient-hospID** of the treated patient's primary hospital through the *putPrivateTreatment* function. The function takes the physician ID, treatment date, hospital ID, and the previously received IPFS link as parameters. We have chosen not to encrypt the private data, but their hash is endorsed, ordered, and committed to the ledger of every peer on the channel as evidence of the transaction for audit purposes. Physicians have the right to write in any of the **Patient-hospID** PDCs, but they can only read from the **Patient-Physician** PDC. To maintain the private data's confidentiality during the transaction flow, we will send them in a transient field in the proposal.

Furthermore, physicians may also record additional data elements in the healthcare record (if necessary), notably medical conditions, family medical history, and surgical details. They will follow the same procedure as the one used when adding a new treatment. However, the family medical history and surgical procedure details will be encrypted using a different symmetric key. Using the same symmetric key will compromise the patient's data since this key is to be shared with the emergency contacts.

The physician's last role consists of updating the emergency details stored on the ledger. These data include patients' medical chronic conditions, current medications, allergies, and adverse drug reactions. Since patient data is rich and distributed across different storage locations, it is crucial to go into the data details before starting the network configuration and chaincode implementation. We must select the function parameters and choose the data format, type, and structure for each data element shared between two distinct network actors while considering our network's data classification and privacy policy. For example, patients' chronic medical conditions are classified as critical data in emergency care. As a

result, when implementing the business logic of the network, we must consider that these data are stored unencrypted on the ledger to assist the emergency physician in providing treatment promptly. Hence, we should keep the minimum useful and non-identifying information on the ledger. When adding a new chronic medical condition, physicians call the *addChroniCondition* function through the physician graphical user interface, which takes the patient emergencyID2 and the International Classification of Diseases (ICD) codes of the diagnosed chronic illness as parameters. The ICD provides a uniform language that allows healthcare professionals to communicate patient information effectively. Besides, physicians use the patient emergencyID1 to retrieve the patient's healthcare record ID and emergencyID2.

5.4.1.1.5 Check-out and billing

We will not discuss this stage in our proposed solution because it is linked to actors not included in the studied ecosystem, notably the health insurance company or the government. It will be part of our future work where we will assess the data classification of healthcare financial activity and discuss data handling after introducing the new actors to our ecosystem.

5.4.1.2 Emergency care

In an emergency care context, we presume the patient to be unconscious and unable to provide the needed information. Otherwise, the situation will be identical to a patient care context, with the patient providing all mandatory data required for treatment. The emergency physician will use the emergencyID1 carried by the patient to acquire, from the **Emergency** PDC, the patient's emergencyID2, blood group, and emergency contact information. Having stored without encryption some data elements from the patient healthcare record on the blockchain ledger, we must maintain the patient's identity hidden from any system intruder. Therefore, we have classified critical data into four categories, each handled and accessed differently:

First category: Data accessible through the emergencyID1 and stored on **Emergency** PDC.

Second category: Data accessible through emergencyID2 and stored on the ledger (public state database). These data include the patient's current medications, medical chronic condition, allergies, and adverse drug reactions.

Third category: Data accessible by the emergency contact or healthcare authority administrator and shared with the emergency physician through the **Patient-Emergency** PDC. The data stored in the **Patient-Emergency** PDC are purged after a specified number of blocks.

- Registration approach 1: Emergency contact shares the emergency administrative and clinical data (family medical history and surgical history) with the emergency physician.
- Registration approach 2: Emergency contact shares the patient's administrative ID and the emergency clinical data with the emergency physician. The emergency physician requests the patient's emergency administration data from the healthcare authority by sharing with them the patient's administrative ID.

Fourth category: Laboratory and diagnoses test are performed immediately on-site once the patient enters the emergency department to enable timely decision-making and intervention. Previous laboratory tests are not mandatory in this case.

Table 5.2: Data distribution and permission rights of network actors across distinct storage locations in our Fabric-based healthcare ecosystems (R: Read, W: Write, NA: No Access)

		Stored Data	Approach 1		Approach 2		Physician	Emergency Physician	Hospital Registrar
			Patient	HA admin	Patient	HA admin			
IPFS	Healthcare Authority (HA)	Patient's demographic and financial information	R	W	NA	R/W	NA	NA	NA

Table 5.2 – *Continued on next page*

Table 5.2 – *Continued from previous page*

	Hospital Organization	De-identified healthcare record (lab tests, surgery details, diagnosis, treatment, and prescriptions, etc.)	R	NA	R	NA	R/W	R/W	NA
Private Data Collections	PatientDemFin	Patient's administrative ID and symmetric keys associated with the IPFS link(s) of the patient's administrative data	R	R/W	NA	R/W	NA	NA	NA
	Patient-hospID	Patient's healthcare record ID, administrative ID, clinical data metadata, and their respective IPFS URL	R	NA	R	NA	W	W	W
	Physician-hospID	Physicians' information: full name, physician ID, medical specialty, contact information, etc.	NA	NA	NA	NA	R/W	R/W	R
	Patient-Physician	Data shared by a patient to a physician after being encrypted with the physician public key	W	NA	W	NA	R	NA	NA

Table 5.2 – *Continued on next page*

Table 5.2 – *Continued from previous page*

	Emergency	Patient's emergencyID1, emergencyID2, gender, blood group, emergency contact information	NA	NA	NA	NA	R	R/W	W
	Patient-Emergency	Data shared by one of patient's emergency contacts (Patient's administrative emergency data and patient's mandatory clinical data)	W	NA	W	W	NA	R	NA
Ledger/public state DB		Patient's emergencyID2, current medications, medical chronic condition, allergies, and adverse drug reactions	R[3]	R[3]	R[3]	R[3]	R/W	R/W	R[3]

5.4.2 Pharma network

The pharma R&D channel represents the second network in our ecosystem that connects a hospital with a pharmaceutical company, notably with its R&D department, to share de-identified and relevant patient data. Pharmaceutical companies are organizations involved in manufacturing, promoting, and distributing drugs and medications. In addition to drug manufacturing, pharmaceutical companies undertake research and development efforts to discover and develop new medicines. The drug development process involves conducting preclinical and clinical research, and obtaining regulatory approvals to ensure the safety and

[3]Even though the patient, healthcare authority administrator and hospital registrar can query the ledger, they don't have the right to read the emergency data stored there.

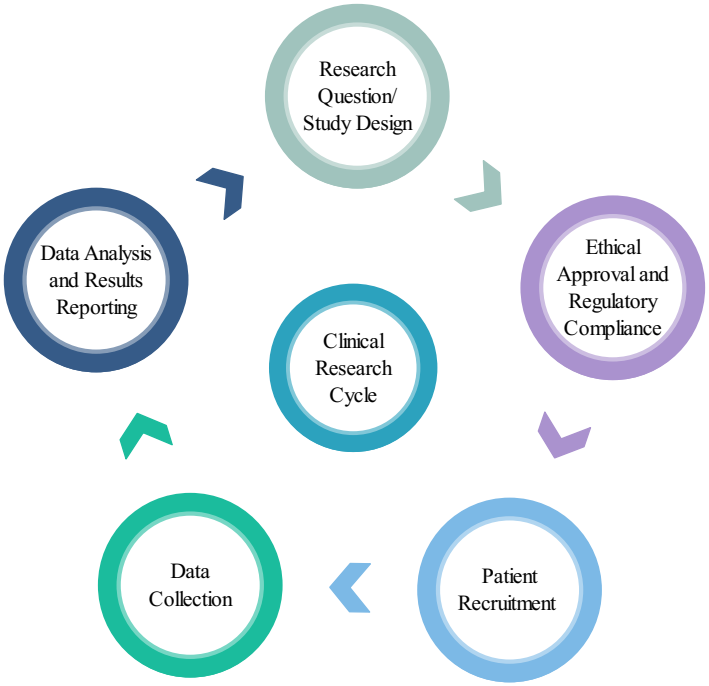


Figure 5.5: Clinical research Cycle [237]

efficacy of the developed drug [236]. While pharmaceutical companies engage in a wide range of activities involving multiple stakeholders, we will focus, in the pharma R&D channel, on the drug development process, notably clinical research. According to the FDA, clinical research refers to studies or trials conducted on people [236]. The different interactions in this network are mainly related to the clinical research cycle depicted in Figure 5.5.

However, to keep it short, our solution will focus on the patient recruitment and data collection phases, representing the interaction between our main healthcare ecosystem actors:

- The clinical researcher affiliated with the R&D department of a pharmaceutical company
- The patient and the physician affiliated with a hospital

As already stated, the physician is taking on the role of both medical practitioner and principal investigator in our use case scenario to narrow our research scope. After defining the research question and having the study plan approved by relevant ethics committees, the clinical researcher can start the clinical trial.

One of the primordial elements of clinical research is recruiting and enrolling qualified participants in clinical trials. Even though traditional advertising and social media marketing are strategies employed to connect with potential participants, recruiting eligible participants within a designated time frame remains a persistent challenge [238]. Besides, engaging with clinical trial ads on social platforms raises concerns about user privacy [239]. To meet these challenges, we need more partnerships with healthcare professionals, such as physicians, and a secure data-sharing environment where patients control how their data is shared and to whom [193], [240]. Physicians are the most efficient at finding and recruiting eligible patients [238], [240], [241]. They can encourage patients to enroll in clinical trials if they have comprehensive details on the clinical trial procedure [241]. It is especially true if the research topic aligns with the physician's areas of interest and is connected to real-world clinical practice [240]. Therefore, as a first step, the clinical researcher will store on the channel ledger all clinical trial details, notably the inclusion and exclusion criteria for ideal participants, the research topic, the study duration. Physicians query the ledger to check the clinical trials currently recruiting patients. They search for the topic that interests them to potentially participate in the related clinical research. Since each channel has its ledger and chaincode, *InvokeChaincode* function is needed to allow physicians to query the pharma R&D channel world state from their graphical user interface. It enables a chaincode to invoke a function in another chaincode in the same or different channel [30]. Each enrolled patient is assigned a research ID. Physicians write the patient's research ID into the patient's healthcare record, namely the **Patient-hospID** PDC of the patient's primary hospital. Patients then use their research ID to collaborate with the clinical researchers after accepting to participate and share their health data. Besides, all the study recruitment details are stored in a private data collection managed by the pharmaceutical company. This PDC contains the study ID associated with the research IDs of the recruited patients, the hospital affiliation of the patients, and the physicians assisting the patients during the clinical trial procedure.

The patient recruitment phase is followed by the data collection phase. Clinical researchers can adopt multiple approaches to collect patient data, notably questionnaire surveys, healthcare records, proxy/informant information, and biological samples [242]. Healthcare records are an important source of high-quality data even though their non-standardized nature can cause some challenges in research analysis. In our use case scenario, patient healthcare records are the first source of information due to their integrity and accuracy since only physicians can write into the patient healthcare record, notably the **Patient-hospID** PDC of the patient's primary hospital. Once patients' consent is obtained, clinical researchers can work

with the appropriate hospitals to gain access to the required healthcare record. As shown in Table 2.1, most data elements are classified as optional. In fact, the data needed by the clinical researcher depend on the research question and on the patients' willingness in participating and sharing their data. The collected data are stored on a repository managed by the pharmaceutical company.

5.4.3 System interaction

This subsection focuses on the interactions among system actors in two different use case scenarios: the registration phase and the consultation phase. The sequence diagrams represent three distinct scenarios; however, we condense these UML diagrams to include only the main components needed to understand our solution functionalities. Figure 5.6 shows the UML sequence diagram of a patient and a physician requesting their identities from their respective organizations. To participate in the Hyperledger Fabric network, users, whether patients or physicians, require a certificate and a private key. After obtaining the enrollment certificate, the user can use the web application to submit a transaction. As mentioned in the patient registration process, the healthcare authority administrator is authorized to write the patient's administrative data to the private IPFS of the healthcare authority (HA-IPFS) and the **PatientDemFin** PDC.

Since we are adopting the first registration approach, the healthcare authority administrator uploads the patient's demographic, financial, and emergency administrative data on the healthcare authority's private IPFS via their graphical user interface (GUI). Three symmetric keys are generated for each patient to encrypt each one of these distinct datasets before being added to the HA-IPFS. The execution of these *addPatientDem*, *addPatientFinInf*, and *addPatientEmgAdminData* functions returns three IPFS hashes (aka URL of the stored files), which are then fed up independently to a chaincode function *putPatient-PrivateData* accessible through the GUI. This function adds these URLs alongside their symmetric key to the **PatientDemFin** PDC under the patient's administrative ID. These keys are encrypted using the patient's public key before storage. In addition, the healthcare authority must share the IPFS URL of the emergency administrative data with the emergency contact. In this case, the symmetric key, used to encrypt the emergency administrative data, is encrypted with the public key of the emergency contact. It is then added alongside the IPFS URL in the **PatientDemFin** PDC under the administrative ID of the patient's

emergency contact.

Aside from the healthcare authority administrator's role in collecting administrative data from patients, the hospital registrar holds responsibilities related to registering new patients. It entails giving each patient a healthcare record ID and instructing the emergency team to assign them emergency IDs. The hospital registration procedure is depicted in Figure 5.7. Figure 5.8 showcases the stages a patient passes through when visiting a physician, with hospital 1 being the patient's primary hospital. It covers stages from the appointment request until the consultation outcome, where the physician enters the information gathered into the patient's medical file. In our use case scenario, the appointment request has occurred outside the blockchain network, and the physician has added a newly diagnosed medical condition. The diagnosis details are encrypted with the patient's symmetric key, previously shared on the **Patient-Physician** PDC, before being added to private IPFS shared among all the hospitals in the network (Hosp-IPFS). After receiving the IPFS URL, the physician must add the diagnosis metadata in the **Patient-hosp1** PDC. The metadata include the physician ID, hospital ID, data domain, diagnosis date, medical condition type (chronic, acute, infectious, etc.), and the IPFS URL. If the newly diagnosed condition is chronic, the physician should record it under the patient's emergencyID2 in the ledger.

5.5 Security validation using AVISPA

In order to evaluate the security of the different interactions in our Fabric-based healthcare ecosystem, we have performed a formal security verification of our proposed protocols through simulation using Automated Validation of Internet Security Protocols and Applications (AVISPA) [243]. AVISPA is a comprehensive toolset for the automated analysis and validation of security protocol written in the High-Level Protocol Specification Language (HLPSL) [244]. It includes four analytic tools (Back-ends), notably OFMC (On-the-Fly Model Checker), CL-AtSe (Constraint-Logic-based At- tack Searcher), SATMC (SAT-based Model-Checker), and TA4SP (Tree Automata-based Protocol Analyzer) [244]. We have used the OFMC and CL-AtSe to analyze if our security goals are satisfied or violated since SATMC and TA4SP are not supported in our adopted SPAN version (1.6) [244].

We have focused our analysis on the following:

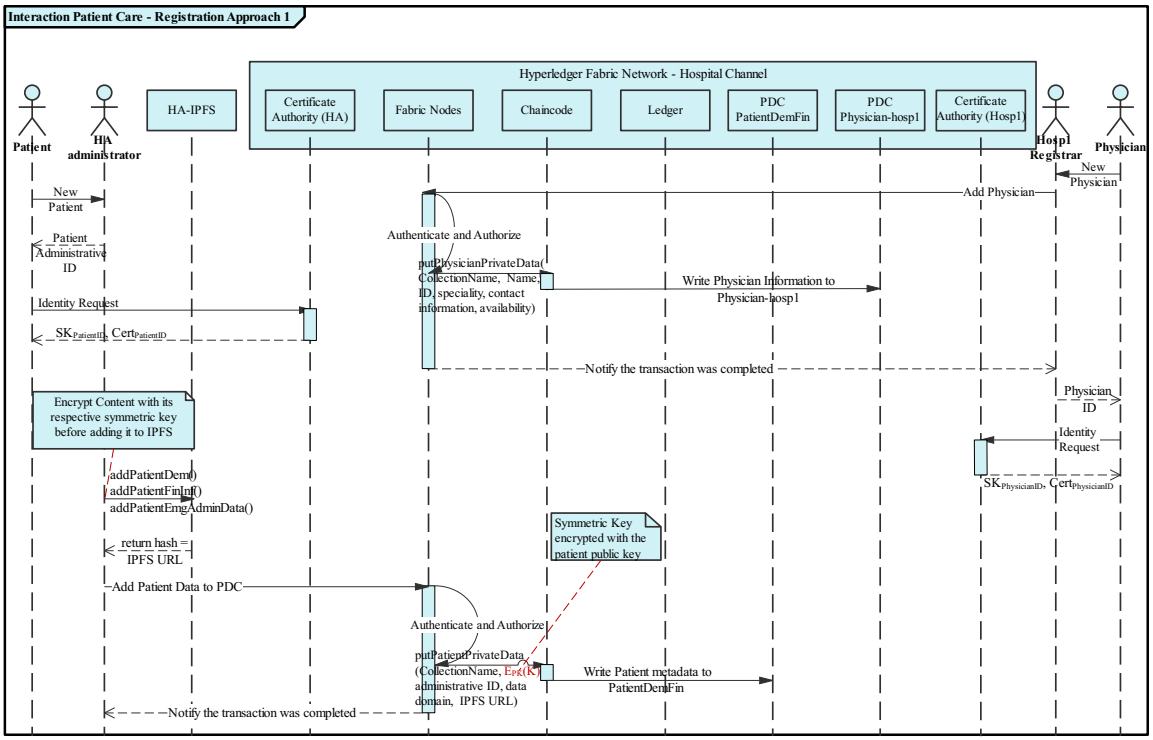


Figure 5.6: UML sequence diagram of registering and enrolling a physician and a patient in the Fabric-based healthcare ecosystem

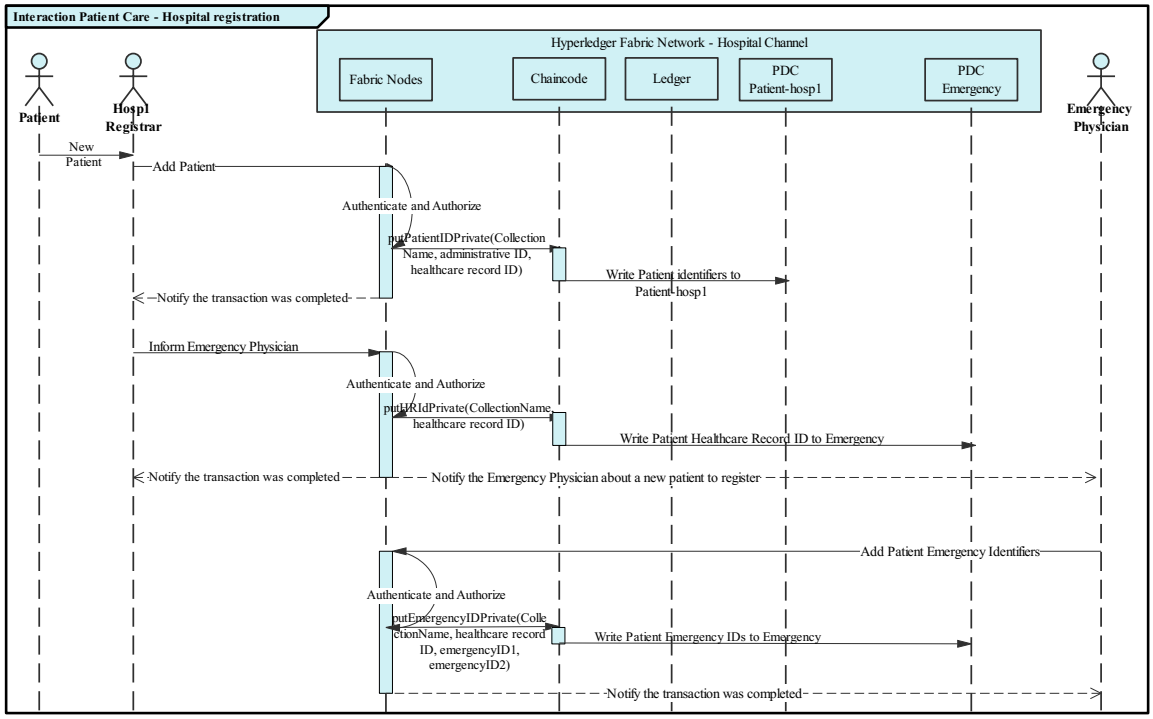


Figure 5.7: UML sequence diagram of the hospital registration procedure

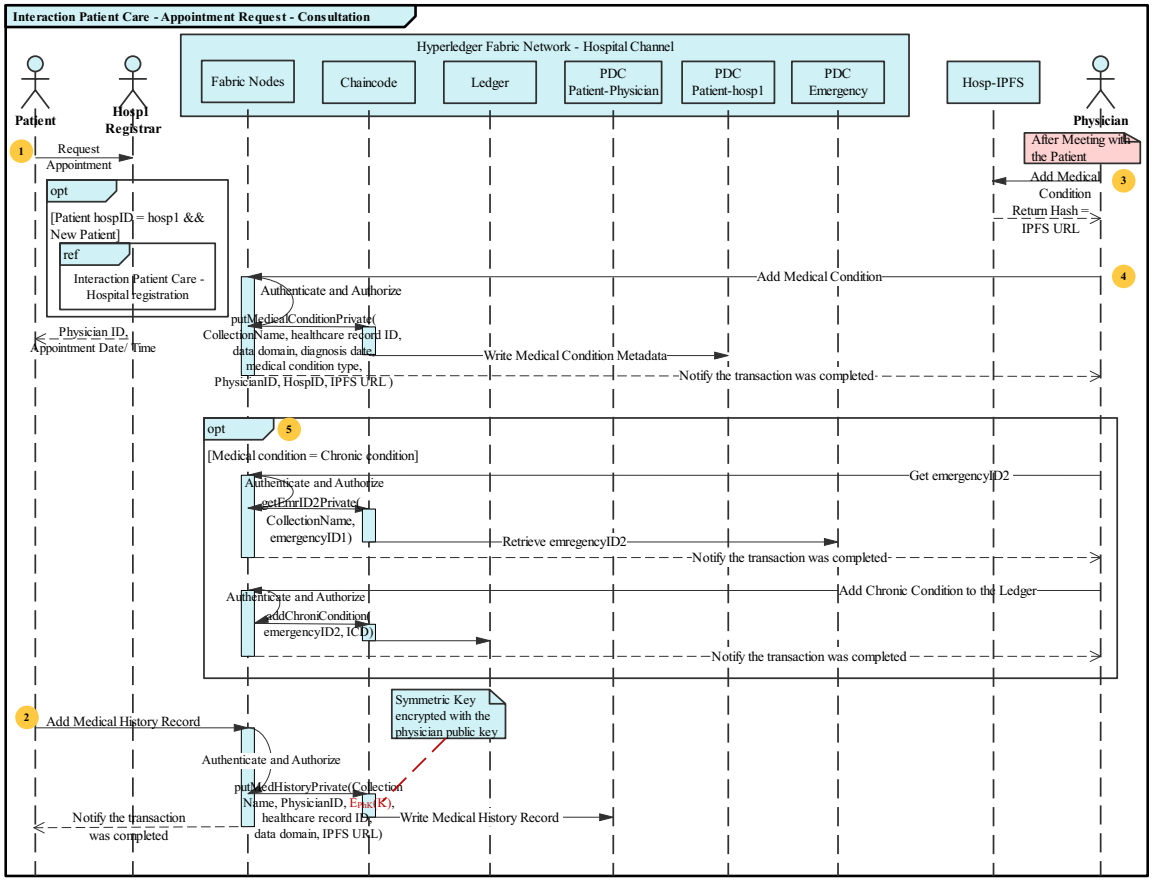


Figure 5.8: UML sequence diagram of the appointment and consultation phases

- Mutual authentication and information secrecy required when exchanging messages and symmetric keys among network participants.

- User authentication and access control authorization required when adding encrypted data to the IPFS or healthcare metadata alongside IPFS URLs to the private data collection.

We have chosen to evaluate these security processes during the registration phase. Because HLPSL is a role-based language, we must define a role for each participant, detailing their activities during the evaluated interaction [244], [245]. We have identified five agent roles in our developed AVISPA model: Patient, HAadmin, Blockchain, IPFS, and PatientDemFinPDC. Table 5.3 lists the various parameters and notations used in the role description. Even though HLPSL has a powerful construct for modeling and verifying security properties, it has certain limitations, particularly in the difficulty of specifying large and complex protocols [245]. Therefore, the evaluation is carried out in three stages to avoid code complexity

and errors in the HLPSL specification and to ensure that each protocol model executes to completion.

Evaluation Phase 1

This phase represents a key exchange protocol between a Patient and an HAadmin through a trusted server (or a secure channel) with mutual authentication of the Patient and HAadmin. We have implemented the role of these three entities in HLPSL language during the registration phase 1 (Figures 5.9,5.10,5.11), where the patient's demographic data and symmetric key are exchanged with secrecy. The healthcare authority administrator initiates the communication by generating a symmetric key and sharing it with the server. The verification results and protocol simulation of the registration phase 1, using OFMC and CL-AtSe, are shown in Figure 5.12. Two secrecy goals and two authentications have been verified.

- *secrecy_of sec_kpd* confirms that the symmetric key Kpd is known only by the patient and healthcare authority administrator.
- *secrecy_of sec_mpd* confirms that the patient's demographic data Mpd is known only by the patient and healthcare authority administrator.
- *authentication_on auth_kpd* confirms that the patient and healthcare authority administrator are using the same key Kpd.
- *authentication_on auth_np* confirms that patient is able to authenticate healthcare authority administrator through Np.

Evaluation Phase 2

This phase represents an interaction between the HAadmin and the IPFS, where only authorized users can add administrative data to the IPFS. We have implemented the role of these two entities in HLPSL language during the registration phase 2 (Figures 5.13, 5.14), where the healthcare authority administrator initiates the communication by sending their user attributes (UA) and certificate (Certad) needed for authentication and access control. The verification results and protocol simulation of the registration phase 2, using OFMC and CL-AtSe, are shown in Figure 5.15. One secrecy goal and two authentications have been verified.

- *secrecy_of sec_empd* confirms that the encrypted demographic data is known only by the healthcare authority administrator and IPFS.

- *authentication_on auth_verif* confirms that IPFS is able to authenticate the healthcare authority administrator through Verify.
- *authentication_on auth_ua* confirms that the healthcare authority administrator is able to authenticate IPFS through UA.

Evaluation Phase 3

This phase represents the interaction among HAadmin, Blockchain, and PatientDemFinPDC. We have implemented the role for these three entities in HLPSL language during the registration phase 3 (Figures 5.16, 5.17, 5.18), where a user authentication and access control authorization check are required to allow a user, in our case HAadmin, to write into the **PatientDemFin** private data collection. Since access control is configured on-chain, at the chaincode level, it is the blockchain's role to assess the user's attribute and, based on the collection policy, grant or deny access to the user to write into the requested PDC. The healthcare authority administrator initiates the communication by an access request, sending their certificate and user attributes to the blockchain. Before granting access authorization, the blockchain authenticates the user and checks that their user attributes comply with the collection policy. We have chosen three different symmetric keys to encrypt the messages exchanged. One of these symmetric keys is owned and known by the HAadmin and the PatientDenFinPDC, noting that the data exchanged is not stored on the blockchain and is only known by the HAadmin and private data collection. The verification results and protocol simulation of the registration phase 3, using OFMC and CL-AtSe, are shown in Figure 5.19. Two secrecy goals and three authentications have been verified.

- *secrecy_of pi* confirms that the collection policy is known only by the blockchain and PatientDemFin private data collection.
- *secrecy_of qm* confirms that the IPFS URL is known only by the healthcare authority administrator and the **PatientDemFin** private data collection.
- *weak_authentication_on auth_access_control* confirms that the blockchain is able to authenticate the healthcare authority administrator through HAadminAtt.
- *authentication_on auth_ck_policy* confirms that the blockchain is able to authenticate the **PatientDemFin** private data collection.
- *authentication_on auth_access* confirms that the blockchain is able to authenticate the healthcare authority administrator through Access.

```
role role_HAadmin(HAadmin,Patient,Server:agent,
                  Khs:symmetric_key,
                  SND,RCV:channel(dy))
played_by HAadmin
def=
    local
         State:nat,
         Np:text,
         Kpd:symmetric_key,
         Mpd:message
    init State:=0
    transition
        1. State=0 /\ RCV(start) =|>
           State':=1 /\ Kpd':=new()
                     /\ SND({Patient.Kpd'}_Khs)
                     /\ secret(Kpd',sec_kpd,{HAadmin,Patient,Server})
        2. State=1 /\ RCV({Patient.Np'}_Kpd) =|>
           State':=2 /\ SND({Np'}_Kpd)
                     /\ request(HAadmin,Patient,auth_kpd,Kpd)
                     /\ witness(HAadmin,Patient,auth_np,Np')
        3. State=2 /\ RCV({Mpd'}_Kpd) =|> State':=3
end role
```

Figure 5.9: HLPSL code for implementing the healthcare authority administrator role during the registration phase 1

```
role role_Server(Server,HAadmin,Patient:agent,
                 Khs,Kps:symmetric_key,
                 SND,RCV:channel(dy))
played_by Server
def=
    local
         State:nat,
         Kpd:symmetric_key
    init State:=0
    transition
        1. State=0 /\ RCV({Patient.Kpd'}_Khs) =|>
           State':=1 /\ SND({HAadmin.Kpd'}_Kps)
end role
```

Figure 5.10: HLPSL code for implementing the trusted server during the registration phase 1

5.6 Discussion

Our proposed Hyperledger Fabric blockchain-based framework facilitates health data exchange among various healthcare system actors while prioritizing patients' privacy. By leveraging blockchain technology, the framework ensures data integrity, immutability, and transparency. Moreover, Hyperledger Fabric operates as a permissioned blockchain with decentralized authority controlling network transactions where all network participants are known to each other, enhancing authentication and instilling trust and security in the system. Nevertheless, these measures are insufficient to address multiple data breaches compromising

Table 5.3: Notations used in the role description

Notation	Description
HAadmin	Healthcare authority administrator
PatientDemFinPDC	private data collection **PatientDemFinPDC**
Certad	Healthcare authority administrator certificate
Np,Verify	Random number
Req	Collection policy request
Pi	Collection policy
Mpd	Patient demographic information
Kpd	Symmetric key used to encrypt the patient demographic information
EMpd	Encrypted patient demographic information
UA	User attributes (such as role/job, organization, department etc.)
Qm	IPFS Hash = IPFS URL
Khs,Kps,Ks1,Kb,Kpdc,Khpdc	Session key

```
role role_Patient(Patient,HAadmin,Server:agent,
                  Kps:symmetric_key,
                  SND,RCV:channel(dy))
played_by Patient
def=
    local
         State:nat,
         Np:text,
         Kpd:symmetric_key,
         Mpd:message
    init State:=0
    transition
        1. State=0 /\ RCV({HAadmin.Kpd'}_Kps) =|>
           State':=1 /\ Np':=new()
                     /\ SND({Patient.Np'}_Kpd')
                     /\ witness(Patient,HAadmin,auth_kpd,Kpd')
        2. State=1 /\ RCV({Np}_Kpd) =|>
           State':=2 /\ request(Patient,HAadmin,auth_np,Np)
                     /\ Mpd':=new()
                     /\ SND({Mpd'}_Kpd)
                     /\ secret(Mpd,sec_mpd,{Patient,HAadmin})
end role
```

Figure 5.11: HLPSL code for implementing the patient role during the registration phase 1

SPAN 1.6 - Protocol Verification : Registration-Phase1.hlpsl

File

```
% OFMC
% Version of 2006/02/13
SUMMARY
  SAFE
DETAILS
  BOUNDED_NUMBER_OF_SESSIONS
PROTOCOL
  /home/span/span/testsuite/results/Registration-Phase1.if
GOAL
  as_specified
BACKEND
  OFMC
COMMENTS
STATISTICS
  parseTime: 0.00s
  searchTime: 71.41s
  visitedNodes: 32400 nodes
  depth: 19 plies
```

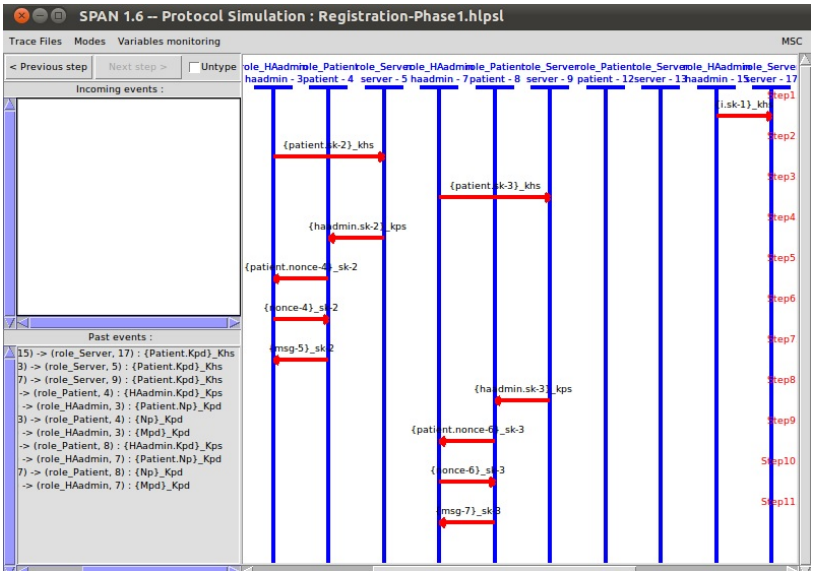
(a) Verification result obtained using OFMC

SPAN 1.6 - Protocol Verification : Registration-Phase1.hlpsl

File

```
SUMMARY
  SAFE

DETAILS
  BOUNDED_NUMBER_OF_SESSIONS
  TYPED_MODEL

PROTOCOL
  /home/span/span/testsuite/results/Registration-Phase1.if

GOAL
  As Specified

BACKEND
  CL-AtSe

STATISTICS

  Analysed   : 29503 states
  Reachable  : 16691 states
  Translation: 0.02 seconds
  Computation: 0.38 seconds
```

(b) Verification result obtained using CL-AtSe



(c) Protocol simulation

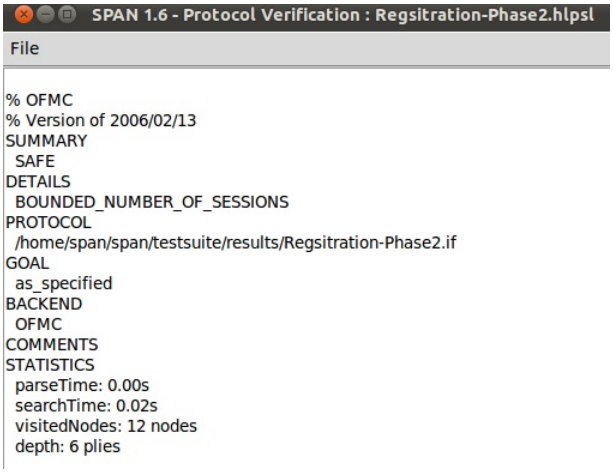
Figure 5.12: Protocol verification and simulation of the registration phase 1

```
role role_HAadmin (HAadmin,IPFS:agent,
                   Certad:text,
                   Ks1:symmetric_key,
                   SND,RCV:channel(dy))
played_by HAadmin
def=
    local
      State:nat,
      UA,Qm,Verify:text,
      EMpd:message
    init State :=0
    transition
    1. State = 0 /\ RCV(start) =|>
       State' := 1 /\ UA':=new()
                   /\ SND({HAadmin.IPFS.Certad.UA'}_Ks1)
    2. State=1 /\ RCV({IPFS.HAadmin.UA.Verify'}_Ks1) =|>
       State':=2 /\ request(HAadmin,IPFS,auth_ua,UA)
                 /\ EMpd':=new()
                 /\ secret(EMpd',sec_empd,{HAadmin,IPFS})
                 /\ SND({Verify'.EMpd'}_Ks1)
                 /\ witness(HAadmin,IPFS,auth_verif,Verify')
    3. State=2 /\ RCV({IPFS.HAadmin.Verify.Qm'}_Ks1) =|> State':=3
end role
```

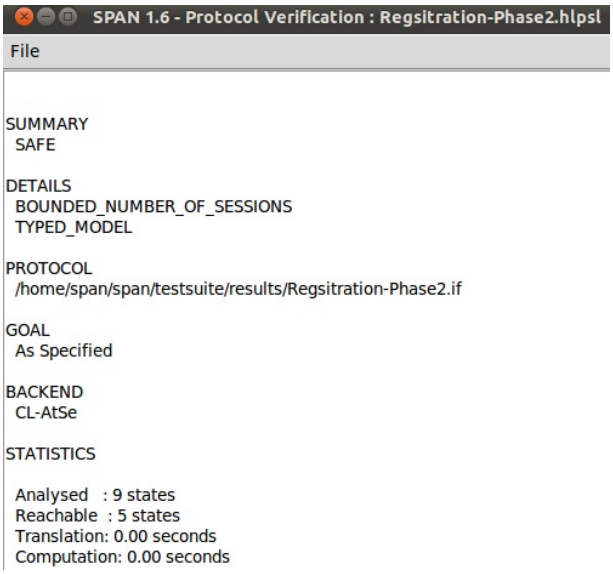
Figure 5.13: HLPSL code for implementing the healthcare authority administrator role during the registration phase 2

```
role role_IPFS (IPFS,HAadmin:agent,
                Certad:text,
                Ks1:symmetric_key,
                SND,RCV:channel(dy))
played_by IPFS
def=
    local
      State:nat,
      EMpd:message,
      Qm,UA,Verify:text
    init State :=0
    transition
    1. State = 0 /\ RCV({HAadmin.IPFS.Certad.UA'}_Ks1) =|>
       State' := 1 /\ Verify':=new()
                   /\ SND({IPFS.HAadmin.UA'.Verify'}_Ks1)
                   /\ witness(IPFS,HAadmin,auth_ua,UA')

    2. State = 1 /\ RCV({Verify.EMpd'}_Ks1) =|>
       State':=2 /\ Qm':=new()
                 /\ SND({IPFS.HAadmin.Verify.Qm'}_Ks1)
                 /\ request(IPFS,HAadmin,auth_verif,Verify)
end role
```

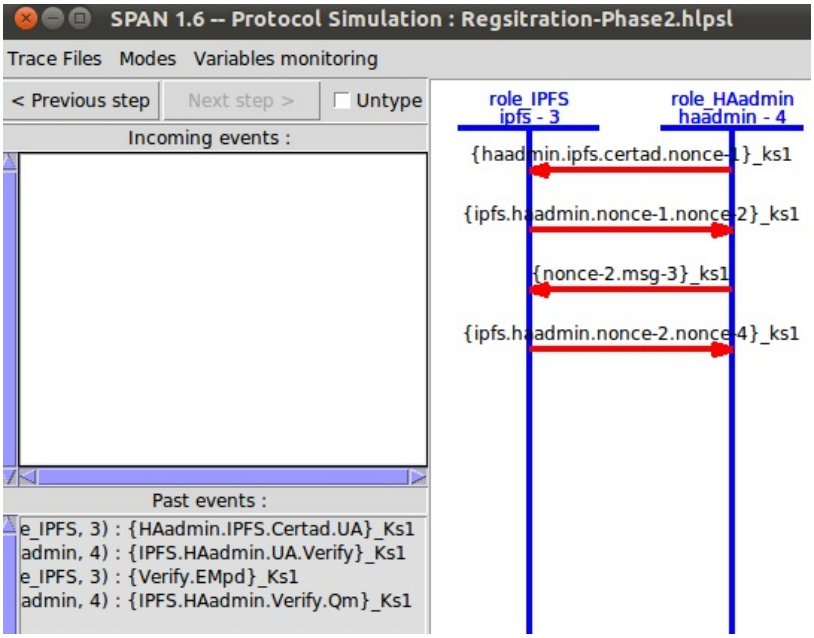
Figure 5.14: HLPSL code for implementing the IPFS role during the registration phase 2



(a) Verification result obtained using OFMC



(b) Verification result obtained using CL-AtSe



(c) Protocol simulation

Figure 5.15: Protocol verification and simulation of the registration phase 2

```
role role_HAadmin(HAadmin,Blockchain,PatientDemFinPDC:agent,
                  Kb,Khpdc:symmetric_key,
                  Certad:text,
                  SND,RCV:channel(dy),
                  Hash: hash_func)
played_by HAadmin
def=
    local
        State: nat,
        UA,Qm,Access:text,
        HAadminAtt:message
    init State :=0
    transition
    1. State = 0 /\ RCV(start) =|>
       State':=1 /\ UA':=new()
                 /\ HAadminAtt':=Hash({HAadmin.Blockchain.Certad.UA'}_Kb)
                 /\ SND({HAadmin.Blockchain.Certad.UA'}_Kb)
                 /\ witness(HAadmin,Blockchain,auth_access_control, HAadminAtt')
    2. State = 1 /\ RCV({Blockchain.HAadmin.Access'}_Kb) =|>
       State':=2 /\ Qm':=new()
                 /\ SND(HAadmin.Blockchain.Access'.Certad.{Qm'}_Khpdc)
                 /\ witness(HAadmin,Blockchain,auth_access,Access')
                 /\ secret(Qm',qm,{HAadmin,PatientDemFinPDC})
end role
```

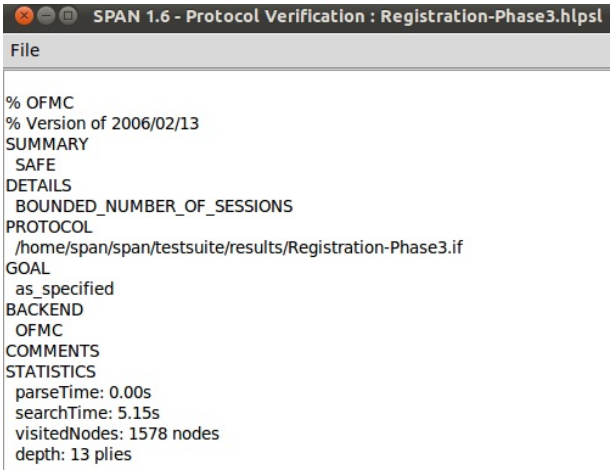
Figure 5.16: HLPSL code for implementing the healthcare authority administrator role during the registration phase 3

```
role role_Blockchain(Blockchain,HAadmin,PatientDemFinPDC:agent,
                  Kb,Kpdc:symmetric_key,
                  Certad:text,
                  SND,RCV:channel(dy),
                  Hash: hash_func)
played_by Blockchain
def=
    local
        State: nat,
        Khpdc:symmetric_key,
        UA,Pi,Qm,Req,Access:text,
        HAadminAtt,HPDCpolicy:message
    init State :=0
    transition
    1. State = 0 /\ RCV({HAadmin.Blockchain.Certad.UA'}_Kb) =|>
       State':=1 /\ HAadminAtt':=Hash({HAadmin.Blockchain.Certad.UA'}_Kb)
                 /\ wrequest(Blockchain,HAadmin,auth_access_control, HAadminAtt')
                 /\ Req':=new()
                 /\ SND({Blockchain.PatientDemFinPDC.Req'}_Kpdc)
    2. State = 1 /\ RCV({PatientDemFinPDC.Blockchain.Pi'}_Kpdc) =|>
       State':=2 /\ HPDCpolicy':=Hash({PatientDemFinPDC.Blockchain.Pi'
                                      }_Kpdc)
                 /\ request(Blockchain,PatientDemFinPDC,auth_ck_policy, HPDCpolicy')
                 /\ Access':=new()
                 /\ SND({Blockchain.HAadmin.Access'}_Kb)
    3. State = 2 /\ RCV(HAadmin.Blockchain.Access.Certad.{Qm'}_Khpdc') =|>
       State':=3 /\ request(Blockchain,HAadmin,auth_access,Access)
                 /\ SND(Blockchain.PatientDemFinPDC.{Qm'}_Khpdc')
end role
```

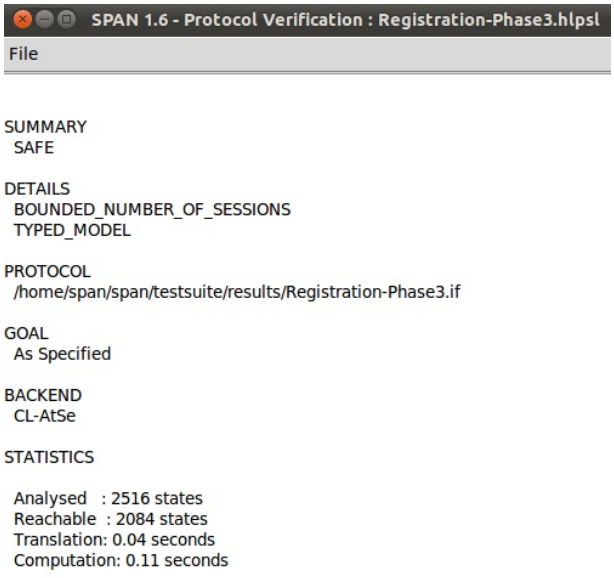
Figure 5.17: HLPSL code for implementing the blockchain role during the registration phase 3

```
role role_PatientDemFinPDC(PatientDemFinPDC,Blockchain:agent,
                           Kpdc,Khpdc:symmetric_key,
                           SND,RCV:channel(dy),
                           Hash: hash_func)
played_by PatientDemFinPDC
def=
    local
        State: nat,
        Pi,Req,Qm:text,
        HPDCpolicy:message
    init State :=0
    transition
    1. State = 0 /\ RCV({Blockchain.PatientDemFinPDC.Req'}_Kpdc) =|>
       State':=1 /\ Pi':=new()
                 /\ HPDCpolicy':=Hash({PatientDemFinPDC.Blockchain.Pi'
                                       }_Kpdc)
                 /\ SND({PatientDemFinPDC.Blockchain.Pi'}_Kpdc)
                 /\ secret(Pi', pi, {PatientDemFinPDC,Blockchain})
                 /\ witness(PatientDemFinPDC, Blockchain, auth_ck_policy, HPDCpolicy')
    2. State = 1 /\ RCV(Blockchain.PatientDemFinPDC.{Qm'}_Khpdc) =|> State':=2
end role
```

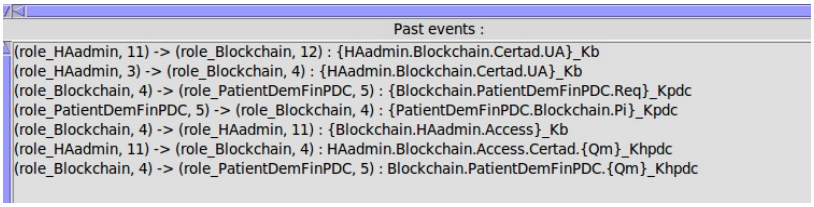
Figure 5.18: HLPSL code for implementing the PatientDemFin PDC role during the registration phase 3



(a) Verification result obtained using OFMC



(b) Verification result obtained using CL-AtSe



(c) Protocol simulation

Figure 5.19: Protocol verification and simulation of the registration phase 3

patient privacy and confidentiality within the healthcare ecosystem. Therefore, our solution introduces data classification and segregation concepts, accompanied by cryptographic and non-cryptographic techniques, to bolster data governance practices concerning data security and privacy.

Data Classification: We have classified healthcare record data based on their relevance in achieving a specific healthcare activity, notably: patient care, emergency care, clinical research, or a healthcare administrative activity. We have started the process by determining the identifiers that should be removed, according to the Safe Harbor de-identification method. We have then classified the data into four categories: critical, mandatory, optional, and restricted. This classification has enabled us to customize the security measures in order to ensure that each participant has access only to the data needed for accomplishing their respective activities.

Data Segregation: Our proposal entails not only segregating data management among various system actors but also segregating classified data within patients' records across separate storage locations. Consequently, no single actor can control all patients' records, enabling granular control over specific data subsets. We have first segregated the healthcare activities by creating two channels in our Fabric network. The hospital channel hosts exchanges related to patient and emergency care activities. The pharma R&D channel hosts exchanges related to clinical research activity. Information privacy and confidentiality are preserved because access to channel resources, notably ledger state, chaincodes, and transactions, is governed by access policies. Secondly, we have segregated data across distinct storage locations: the blockchain ledger (state database), IPFS, and private data collection. Using multiple storage locations has helped implement fine-grained access control, reducing the risk of unauthorized data exposure. Two private IPFSs have been deployed, storing patients' administrative data and patients' de-identified clinical data, respectively. The private data concept introduced by Hyperledger Fabric enables secure data sharing among selected participants within the same channel. Private data are stored on a private state database (SideDB), and the hash of these data is recorded on the public ledger. The latter is used as transaction proof for audit purposes. We have introduced multiple private data collections into our ecosystem, notably:

- **PatientDemFin** PDC, owned and managed by the healthcare authority. It contains patients' admin-

istrative metadata.

- **Patient-hospID** PDC and **Physician-hospID** PDC, owned and managed by each hospital independently from the other. **Patient-hospID** contains the clinical metadata of patients affiliated with hospID. **Physician-hospID** contains information on physicians affiliated with hospID.
- **Patient-Physician**, **Emergency**, and **Patient-Emergency** are three PDCs shared among all hospitals.

Additionally, data segregation has played a vital role in reducing the risk of patient identification through network monitoring. Splitting patient data across multiple locations and assigning different IDs to the same patients in various locations obfuscate the association among different data fragments. For example, patients' emergency data were split between the ledger and the **Emergency** PDC. An emergencyID1 was associated with the data stored on the **Emergency** PDC and an emergencyID2 was associated with the data stored on the ledger. This approach makes it more challenging for unauthorized individuals to correlate and assemble complete patient profiles.

De-identification: We have not used patients' and physicians' real-life identities when enrolling them on the Fabric network. For the patients, we have selected either to use their administrative ID or their healthcare record ID, depending on the adopted approach. Therefore, patients cannot be identified or linked to their real-life identity from their network interactions. As for the physician, we have used their physician ID. These IDs are recognized only by authorized network actors. In addition, we have chosen to store on the IPFS shared among all network hospitals de-identified patients' clinical data, as per Safe Harbor de-identification method.

Access Control: We have implemented attribute-based access policies into the chaincode logic to manage access to the private data collections, IPFSs, and channels' resources. Furthermore, we have empowered patients with control over their data by enabling them to choose whether to share the mandatory or optional data with their preferred physician. Patients also have the responsibility of sharing their de-identified healthcare data with research and third-party organizations, particularly if they choose to participate in a clinical trial, for instance.

Data Encryption: We have encrypted all the data stored on IPFS with the patient's symmetric key. Data encryption helps preserve the confidentiality of data. Multiple symmetric keys have been employed to protect patient data, particularly since we need to share some information with the patient's emergency contact. The symmetric key used to encrypt administrative data differs from the one used to encrypt clinical or emergency data. Using a single symmetric key to access all patient data could compromise patients' privacy and increase the risk of unauthorized access.

HIPAA and GDPR compliance: Our framework adheres to the regulations of HIPAA and GDPR to ensure compliance with data privacy and security standards. Patients maintain control over their data, and only de-identified information is shared with clinical researchers, subject to patient approval. Additionally, under certain conditions, it is possible to unpin data stored on IPFS and delete data stored on private data collection, granting patients the right to erasure. Although patients cannot directly modify data stored on the network, they possess the capability to notify authorized actors to make required corrections.

Moreover, we have presented a solution in which the physician acts as a vital intermediary between the clinical researcher and the patient. The physician identifies patients who meet the researcher's inclusion criteria. Subsequently, the decision to participate in the clinical trial and share de-identified data lies with the patient. This approach streamlines the development of pertinent clinical research and paves the way for evidence-based treatments to be incorporated into routine clinical services.

5.6.1 A generalized privacy-preserving Fabric-based data governance framework

Finally and most importantly, we must highlight the fact that the privacy-preserving Fabric-based data governance framework, we propose in this work, has been designed to be generalizable and seamlessly customizable in order to enable its adoption in any data-driven domain. This can simply be achieved by following these methodological steps:

- Identify the domain stakeholders, their respective roles, and activities.
- Identify the types of data relevant to the designated domain of knowledge and enumerate the various data elements within each type. Particularly focus on breaking down data elements, especially when they encompass sensitive sub-elements such as dates or personally identifiable information.
- Classify data based on their intended use by assigning each type of data and each data element its level of relevance for a specific activity.
- Identify the regulations and guidelines applicable in the designated domain and evaluate their impact on the data elements. Assess whether specific data elements are restricted to certain activities or if they qualify as personally identifiable information, requiring meticulous handling and safeguarding.
- Select the off-chain storage suitable for the developed framework. The off-chain storage is used to store large volumes of data without any personal identification information.
- Establish different channels to partition the network, allowing for private and isolated communication pathways among specific participants, if required. Several factors can impact such a decision, notably:
  - Data and information flow: Data stored on the channel ledger must be relevant to the workflow.
  - Legal, regulatory, and compliance considerations about data shared or transacted in the network.
- Represent the channel activity as a workflow for an efficient data handling, especially in data segregation and access control.
- For each channel:
  - Identify the organizations participating in the channel and the network architecture (number of peers, ordering service, MSPs, CAs, etc.)
  - Identify the number of private data collections to deploy. Define for each collection its properties (name, policy, blockToLive, memberOnlyRead, memberOnlyWrite, etc.). Make sure to deploy a small number of PDCs for more management efficiency.
  - Segregate the identified data elements on the different storage locations: PDCs, ledger, and off-chain storage.

  - Develop and deploy the chaincode. In this part, it is critical to identify the function parameters and choose the data type and structure for each data element shared among the system actors based on the data classification and privacy rules. Data can be summarized in a table and accessed by a group of people, while a more detailed document can be restricted to all network participants except its owner.
  - Implement access control at the chaincode level, taking into consideration the enforced regulations and data classification.

- Identify the cryptographic techniques (encryption, hashing, etc.) that must be introduced to add more confidentiality and privacy.

5.7 Conclusion

Data confidentiality is one of the biggest challenges faced by organizations since data theft and their use for criminal purposes are common. In this work, we address the critical issues pertaining to data confidentiality and privacy within the healthcare ecosystem. We implement a Hyperledger Fabric-based healthcare ecosystem where we deploy a granular access control to protect patients' data from being exposed to unauthorized parties or tracked and traded without the patient's consent. In addition to cryptographic and non-cryptographic techniques, we adopt data classification and segregation concepts to manage and secure data exchange more efficiently while complying with HIPAA and GDPR. Compared to solutions found in the literature, our framework gives individuals the power to erase previously added data. It also provides them with more control over their medical data as they get to choose what to share and with whom. Furthermore, our framework ensures the de-identification of all data being shared for research purposes, following the Safe Harbor de-identification method. Data segregation helps reduce the risk of exposure and potential misusage of patient data. The segregation of patient data across various locations, each accessible by specific authorized personnel, combined with the adoption of multiple identifiers, strengthens the protection of the patient's overall privacy. It becomes challenging for network actors to link a patient's real-life identity to their healthcare record. Besides, our framework provides clinical researcher at pharmaceutical company legitimate medical data and helps them save time in data curation by reducing the expenditure on fake data. Furthermore, and most interestingly, what we have developed is a generalized

approach derived from our framework methodology, enabling its applicability to any domain, extending beyond healthcare exclusively. Knowing that the data market, in general, is very lucrative, the usage of our blockchain-based data governance framework could open the way to all kinds of safe data monetization applications. In the specific context of the healthcare industry, patients would be able to make profits from willingly selling and sharing their health records with third parties such as pharmaceutical companies. However, monetizing patient's health data is another challenge that should be addressed since it requires patient identification before conducting financial transactions with them. The potential of monetizing patients' health data resides in maintaining a private environment that allows them to conduct financial transactions without compromising their privacy. A more enhanced framework would be needed to leverage the use of the blockchain and create that private environment.

Chapter 6

Conclusion and Future Work

This thesis presents a rigorous approach to integrating and implementing blockchain technology in the healthcare industry. We have addressed three challenges that affect the safety and/or privacy of patients.

In the first challenge, we aimed to enhance the drug journey from the manufacturer to the customer. Our goal was to elaborate on the different requirements to ensure that only high-quality and authentic drugs reach the customer and to improve supply chain visibility and management. Our theoretical and real-world application studies have showed how the blockchain reduces risks associated with tracking systems and data management. The blockchain improves chain transparency, enabling fraud prevention, product origin tracking, and efficient product recalls. However, there are key considerations for building a blockchain-based supply chain. We cite the blockchain selection based on ecosystem requirements such as performance or privacy needs. We also must consider the dual storage architecture we need to implement and choose the tracking device and the communication protocol based on the criteria required by the supply chain environment. Besides, device authentication and data encryption are a must to establish a secure tracking system.

In the second challenge, we aimed to improve the patient journey by enhancing the management of patient data and the interaction between physicians and patients. We specifically targeted the management of patient drug prescriptions and drug allergies since adverse drug events are estimated to be within the top

ten of the most common causes of death worldwide. We designed a Hyperledger Fabric-based framework that provides a secure infrastructure for data sharing while giving patients control over who can access their data. The system prioritizes patient privacy and complies with HIPAA regulations through security measures, including access control and data encryption. We implemented private data collections, a Fabric feature, to enable selective data sharing among network participants, enhancing privacy within the network. We also introduced, into our ecosystem, a graph database as an off-chain storage to store our defined drug ontology. Hence, we stored the URI of the prescribed drugs on the blockchain, which not only reduced the strain on the blockchain storage capacity but also enhanced data accessibility. URIs serve as pointers, allowing us to retrieve associated information when needed without burdening the blockchain with the full data load. Besides, using a graph database as off-chain storage reduces query response time. Additionally, the semantic description of the drugs and their composition, as well as all relevant information, enabled knowledge inference from a few given facts. The system would notify the physician of any potential drug-drug interaction or drug-allergy reaction based on two factors, notably the patient's allergies and current medications.

In the third challenge, we aimed to enhance the confidentiality and privacy of healthcare data when exchanged among healthcare actors. We presented a novel approach to data governance by implementing a Hyperledger Fabric-based framework that enables health data exchange among various healthcare system actors while maintaining patient privacy. Our proposed framework is HIPAA and GDPR-compliant, ensuring alignment with the patient data privacy rules. We adopted a comprehensive data security strategy that includes data classification and data segregation to enhance access control management, a key part of data governance. The data classification process involved health data assessment based on its relevance and criticality in achieving one of the studied healthcare activities. On the other hand, the data segregation process involved separating collected health data across various storage locations (off or on-chain) based on the data classification scheme and security requirements. Each storage location consisted of subsets of patients' data, each subset associated with one of the patient's distinct IDs. We used multiple identities for a single patient to isolate specific data and avoid its association with other data fragments, enhancing patient privacy. Additionally, we implemented cryptographic and non-cryptographic measures such as access controls, encryption, and de-identification to improve the security of our data. We took advantage of the channel and private data features of the Hyperledger Fabric to enhance confidentiality. Our proposed solution was then mapped into a more generalized approach to make it applicable in different contexts and

industries.

6.1 Future Work

Although this work provided a sturdy basis for unlocking the potential of blockchain technology and addressing its challenges when adopted in the healthcare industry, there are some unaddressed challenges. As previously defined, the healthcare ecosystem consists of multiple actors who coexist and interact with each other. It comprises healthcare organizations, medical equipment suppliers, healthcare providers, the government, insurance companies, pharmaceutical companies, R&D labs, and patients. Thus, it is highly unlikely for all healthcare actors with millions of users and different infrastructures to maintain the same decentralized blockchain architecture. Therefore, the healthcare ecosystem involves multiple blockchain networks that use different blockchain technologies with different architectures and consensus protocols. These networks must interact and communicate with each other. Cross-blockchain is essential for promoting collaboration among different blockchain networks and achieving interoperability, making the decentralized ecosystem more resilient and adaptable. Thus, for future work, it is interesting to address the challenges associated with cross-blockchain technology, notably the challenge of designing an interoperable system that finds a balance between security and governance.

On a different level, we must acknowledge the growth of the wearable technology market across industries and the integration of artificial intelligence (AI) in wearable devices to provide personalized insights and services. Consumers in the healthcare industry have access to various devices that offer different functionalities. The adoption of such technologies is a game changer, enticing individuals to adopt them with the promise of improving their health care and wellness but at the expense of their privacy. The increasing volume of data puts personal privacy at greater risk. In this thesis, we have focused on addressing the challenges related to healthcare record management. While these data are sensitive and highly valuable, the data trajectory and primary actors' role are clear, allowing us to contain risks to a certain degree. In contrast, when we venture into the realm of wearable devices and AI, the dynamics change significantly, and the tracking of data cannot be easily predicted. Thus, for future work, it is interesting to address the challenge of finding the right balance between using privacy and data as a currency to improve

user experience.

Publications

Journal

- Azzi R, Chamoun RK, Sokhn M. The power of a blockchain-based supply chain. Computers & industrial engineering. 2019 Sep 1;135:582-92.

Conference

- Azzi R, Kilany Chamoun R, Serhrouchni A, Sokhn M. A Healthcare Delivery System Powered by Semantic Data Description and Blockchain. InFuture of Information and Communication Conference 2023 Feb 27 (pp. 224-242). Cham: Springer Nature Switzerland.

References

[1] S. P. Murphy, *Healthcare Information Security and Privacy*. McGraw-Hill Education, Jan. 5, 2015, 560 pp., Google-Books-ID: t5XtoAEACAAJ, ISBN: 978-0-07-183179-6.

[2] S. Kitsiou, A. Matopoulos, V. Manthou, and M. Vlachopoulou, "Evaluation of integration technology approaches in the healthcare supply chain", *International Journal of Value Chain Management*, vol. 1, no. 4, pp. 325–343, Jan. 2007, Publisher: Inderscience Publishers, ISSN: 1741-5357. DOI: 10.1504/IJVCM.2007.015091. [Online]. Available: https://www.inderscienceonline.com/doi/abs/10.1504/IJVCM.2007.015091 (visited on 02/03/2023).

[3] R. Beech, *Health Operations Management: Patient Flow Logistics in Health Care*. Psychology Press, 2005, 345 pp., Google-Books-ID: 4fYNr091F_4C, ISBN: 978-0-415-32396-3.

[4] "Healthcare data breach statistics", HIPAA Journal. (), [Online]. Available: https://www.hipaajournal.com/healthcare-data-breach-statistics/ (visited on 02/04/2023).

[5] "Centralized vs. decentralized data systems —which choice is best?", VentureBeat. (Sep. 12, 2022), [Online]. Available: https://venturebeat.com/data-infrastructure/centralized-vs-decentralized-data-systems-which-choice-is-best/ (visited on 02/04/2023).

[6] S. Goosen. "Data: The public health supply chain's biggest challenge and opportunity", PFSCM. (), [Online]. Available: https://pfscm.org/pfscm_news/data-the-public-health-supply-chains-biggest-challenge-and-opportunity/ (visited on 02/04/2023).

[7] S. Nakamoto, "Bitcoin: A peer-to-peer electronic cash system",

[8] S. Underwood, "Blockchain beyond bitcoin", *Communications of the ACM*, vol. 59, no. 11, pp. 15–17, Oct. 28, 2016, ISSN: 0001-0782. DOI: 10.1145/2994581. [Online]. Available: https://doi.org/10.1145/2994581 (visited on 02/05/2023).

[9] J. Muckstadt, D. Murray, J. Rappold, and D. Collins, "Guidelines for collaborative supply chain system design and operation", *Information Systems Frontiers*, vol. 3, no. 4, pp. 427–453, 2001, ISSN: 1387-3326. DOI: 10.1023/A:1012824820895.

[10] N. T. Government. "Consumer rights". Last Modified: 2019-06-04T18:41:50+09:30 Publisher: https://consumeraffairs.nt.gov.au/for-consumers/consumer-rights. (Apr. 1, 2019), [Online]. Available: https://consumeraffairs.nt.gov.au/for-consumers/consumer-rights (visited on 10/08/2022).

[11] O. o. t. Commissioner. "What we do", FDA. Publisher: FDA. (Jun. 28, 2021), [Online]. Available: https://www.fda.gov/about-fda/what-we-do (visited on 02/08/2023).

[12] O. o. R. Affairs. "Recalls, market withdrawals, & safety alerts", FDA. Publisher: FDA. (Nov. 2, 2022), [Online]. Available: https://www.fda.gov/safety/recalls-market-withdrawals-safety-alerts (visited on 02/07/2023).

[13] N. Kshetri, "1 blockchain's roles in meeting key supply chain management objectives", *International Journal of Information Management*, vol. 39, pp. 80–89, Apr. 1, 2018, ISSN: 0268-4012. DOI: 10.1016/j.ijinfomgt.2017.12.005. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0268401217305248 (visited on 06/27/2020).

[14] S. A. Abeyratne and R. P. Monfared, "Blockchain ready manufacturing supply chain using distributed ledger", *International Journal of Research in Engineering and Technology*, vol. 05, no. 9, pp. 1–10, Sep. 25, 2016, ISSN: 23217308, 23191163. DOI: 10.15623/ijret.2016.0509001. [Online]. Available: https://ijret.org/volumes/2016v05/i09/IJRET20160509001.pdf (visited on 06/15/2020).

[15] K. Toyoda, P. T. Mathiopoulos, I. Sasase, and T. Ohtsuki, "A novel blockchain-based product ownership management system (POMS) for anti-counterfeits in the post supply chain", *IEEE Access*, vol. 5, pp. 17 465–17 477, 2017, Conference Name: IEEE Access, ISSN: 2169-3536. DOI: 10.1109/ACCESS.2017.2720760.

[16] B. Kevin, *ERP : Régler les 10 problèmes de cybersécurité les plus courants*, fr, Nov. 2021. [Online]. Available: https://www.lemagit.fr/conseil/ERP-reglez-les-10-defis-de-securites-les-plus-courants (visited on 03/17/2024).

[17] S. Laaper, W. Yeh, J. Fitzgerald, M. Basir, and E. Quasney. "Using blockchain to drive supply chain transparency and innovation", Deloitte United States. Library Catalog: www2.deloitte.com. (), [Online]. Available: https://www2.deloitte.com/us/en/pages/operations/articles/blockchain-supply-chain-innovation.html (visited on 06/27/2020).

[18] K. M. Eisenhardt, "Building theories from case study research", *The Academy of Management Review*, vol. 14, no. 4, pp. 532–550, 1989, Publisher: Academy of Management, ISSN: 0363-7425. DOI: 10.2307/258557. [Online]. Available: https://www.jstor.org/stable/258557 (visited on 02/08/2023).

[19] M. Kirejczyk, A. Kędracki, I. Rukhavets, and V. Trifa. "Ambrosus whitepaper", The Whitepaper Database. (May 2, 2018), [Online]. Available: https://www.allcryptowhitepapers.com/ambrosus-whitepaper/ (visited on 02/08/2023).

[20] *Modum white paper data integrity for supply chain operations powered by blockchain technology*, 2017. [Online]. Available: https://assets.modum.io/wp-content/uploads/2017/08/modum-whitepaper-v.-1.0.pdf.

[21] Walid Ammar MD, Ph.D. "Health beyond politics-2009". (), [Online]. Available: http://www.moph.gov.lb/en/DynamicPages/view/3908/health-beyond-politics (visited on 01/10/2021).

[22] W. v. Lerberghe, A. Mechbal, N. Kronfol, and Ministry of Public Health, *The Collaborative Governance of Lebanon's Health Sector: Twenty Years of Efforts to Transform Health System Performance*. Republic of Lebanon Ministry of Public Health, 2018, OCLC: 1136101968, ISBN: 978-9953-0-4542-9.

[23] D. W. Bates, D. L. Boyle, M. B. V. Vliet, J. Schneider, and L. Leape, "Relationship between medication errors and adverse drug events", *Journal of General Internal Medicine*, vol. 10, no. 4, pp. 199–205, Apr. 1, 1995, ISSN: 1525-1497. DOI: 10.1007/BF02600255. [Online]. Available: https://doi.org/10.1007/BF02600255 (visited on 05/04/2022).

[24] E. Ammenwerth, P. Schnell-Inderst, C. Machan, and U. Siebert, "The effect of electronic prescribing on medication errors and adverse drug events: A systematic review", *Journal of the American Medical Informatics Association*, vol. 15, no. 5, pp. 585–600, Sep. 1, 2008, ISSN: 1067-5027. DOI: 10.1197/jamia.M2667. [Online]. Available: https://doi.org/10.1197/jamia.M2667 (visited on 05/04/2022).

[25] A. E. Lanham, G. Cochran, and D. Klepser, "Electronic prescriptions: Opportunities and challenges for the patient and pharmacist", 2016. DOI: 10.2147/AHCT.S64477.

[26] B. Aldughayfiq and S. Sampalli, "Digital health in physicians' and pharmacists' office: A comparative study of e-prescription systems' architecture and digital security in eight countries", *Omics: A Journal of Integrative Biology*, vol. 25, no. 2, pp. 102–122, Feb. 2021, ISSN: 1557-8100. DOI: 10.1089/omi.2020.0085.

[27] P. M. P. Office, *HIPAA : Avis sur les procédures de confidentialité - Juin 2016*, Fr, Jun. 2016. [Online]. Available: https://www.pennmedicine.org/for-patients-and-visitors/patient-information/hipaa-and-privacy/~/media/e0235df58bc74cf5b6726c3cd1a91b17.ashx (visited on 03/17/2024).

[28] D. Yaga, P. Mell, N. Roby, and K. Scarfone, "Blockchain technology overview", National Institute of Standards and Technology, Gaithersburg, MD, NIST Internal or Interagency Report (NISTIR) 8202, Oct. 2018, NIST IR 8202. DOI: 10.6028/NIST.IR.8202. [Online]. Available: https://nvlpubs.nist.gov/nistpubs/ir/2018/NIST.IR.8202.pdf (visited on 12/06/2020).

[29] "GraphDB downloads and resources". (), [Online]. Available: https://graphdb.ontotext.com/ (visited on 10/27/2020).

[30] "Key concepts — hyperledger-fabricdocs master documentation". (), [Online]. Available: https://hyperledger-fabric.readthedocs.io/en/release-1.4/key_concepts.html (visited on 12/17/2020).

[31] A. Tanner, *Our Bodies, Our Data: How Companies Make Billions Selling Our Medical Records*. Beacon Press, Jan. 10, 2017, 226 pp., Google-Books-ID: e_WxDQAAQBAJ, ISBN: 978-0-8070-3334-0.

[32] "Medical-record software companies are selling your health data", thestar.com. Section: Investigations. (Feb. 20, 2019), [Online]. Available: https://www.thestar.com/news/investigations/2019/02/20/medical-record-software-companies-are-selling-your-health-data.html (visited on 11/06/2022).

[33] "How can patients make money off their medical data?" (Jan. 29, 2019), [Online]. Available: https://news.bloomberglaw.com/pharma-and-life-sciences/how-can-patients-make-money-off-their-medical-data (visited on 10/23/2022).

[34] T. Helm and T. H. P. Editor, "Revealed: How drugs giants can access your health records", *The Observer*, Feb. 8, 2020, ISSN: 0029-7712. [Online]. Available: https://www.theguardian.com/technology/2020/feb/08/fears-over-sale-anonymous-nhs-patient-data (visited on 11/06/2022).

[35] H. Journal. "Study explores how medical apps are sending health data to facebook and others", HIPAA Journal. (Aug. 26, 2022), [Online]. Available: https://www.hipaajournal.com/study-explores-how-medical-apps-are-sending-health-data-to-facebook-and-others/ (visited on 11/07/2022).

[36] C. Ross. "Call it data liberation day: Patients can now access all their health records digitally", STAT. (Oct. 6, 2022), [Online]. Available: https://www.statnews.com/2022/10/06/health-data-information-blocking-records/ (visited on 12/28/2022).

[37] H. Journal. "Ransomware attacks drop by 23% globally but increase by 328% in healthcare", HIPAA Journal. (Jul. 29, 2022), [Online]. Available: https://www.hipaajournal.com/ransomware-attacks-drop-by-23-globally-but-increase-by-328-in-healthcare/ (visited on 12/31/2022).

[38] S. Shi, D. He, L. Li, N. Kumar, M. K. Khan, and K.-K. R. Choo, "Applications of blockchain in ensuring the security and privacy of electronic health record systems: A survey", *Computers & Security*, vol. 97, p. 101 966, Oct. 1, 2020, ISSN: 0167-4048. DOI: 10.1016/j.cose.2020.101966. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S016740482030239X (visited on 11/22/2021).

[39] L. D. Costa, B. Pinheiro, W. Cordeiro, R. Araújo, and A. Abelém, "Sec-health: A blockchain-based protocol for securing health records", *IEEE Access*, vol. 11, pp. 16 605–16 620, 2023, Conference Name: IEEE Access, ISSN: 2169-3536. DOI: 10.1109/ACCESS.2023.3245046.

[40] B. Zaabar, O. Cheikhrouhou, F. Jamil, M. Ammi, and M. Abid, "HealthBlock: A secure blockchain-based healthcare data management system", *Computer Networks*, vol. 200, p. 108 500, Dec. 9, 2021, ISSN: 1389-1286. DOI: 10.1016/j.comnet.2021.108500. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1389128621004382 (visited on 03/26/2023).

[41] deloitteeditor. "Managing data risk: A new strategic imperative", WSJ. Section: CIO Journal. (), [Online]. Available: https://deloitte.wsj.com/cio/2019/04/04/managing-data-risk-a-new-strategic-imperative/ (visited on 02/03/2023).

[42] M. Gupta, "Blockchain for dummies®, 2nd IBM limited edition", 2018.

[43] T. T. A. Dinh, J. Wang, G. Chen, R. Liu, B. C. Ooi, and K.-L. Tan, "BLOCKBENCH: A framework for analyzing private blockchains", in *Proceedings of the 2017 ACM International Conference on Management of Data*, ser. SIGMOD '17, New York, NY, USA: Association for Computing Machinery, May 9, 2017, pp. 1085–1100, ISBN: 978-1-4503-4197-4. DOI: 10.1145/3035918.3064033. [Online]. Available: https://doi.org/10.1145/3035918.3064033 (visited on 09/30/2020).

[44] Z. Zheng, S. Xie, H.-N. Dai, X. Chen, and H. Wang, "Blockchain challenges and opportunities: A survey", *International Journal of Web and Grid Services*, vol. 14, no. 4, pp. 352–375, Jan. 2018, Publisher: Inderscience Publishers, ISSN: 1741-1106. DOI: 10.1504/IJWGS.2018.095647. [Online]. Available: https://www.inderscienceonline.com/doi/abs/10.1504/IJWGS.2018.095647 (visited on 02/06/2023).

[45] "A blockchain platform for the enterprise — hyperledger-fabricdocs main documentation". (), [Online]. Available: https://hyperledger-fabric.readthedocs.io/en/release-2.2/index.html (visited on 06/13/2022).

[46] L. S. Sankar, M. Sindhu, and M. Sethumadhavan, "Survey of consensus protocols on blockchain applications", in *2017 4th International Conference on Advanced Computing and Communication Systems (ICACCS)*, Jan. 2017, pp. 1–5. DOI: 10.1109/ICACCS.2017.8014672.

[47] T. Bocek and B. Stiller, "Smart contracts – blockchains in the wings", in *Digital Marketplaces Unleashed*, C. Linnhoff-Popien, R. Schneider, and M. Zaddach, Eds., Berlin, Heidelberg: Springer, 2018, pp. 169–184, ISBN: 978-3-662-49275-8. DOI: 10.1007/978-3-662-49275-8_19. [Online]. Available: https://doi.org/10.1007/978-3-662-49275-8_19 (visited on 02/06/2023).

[48] S. McNew. "Council post: How blockchain can solve today's medical supply chain flaws and improve responses for future crises", Forbes. Section: Leadership. (), [Online]. Available: https://www.forbes.com/sites/forbesbusinessdevelopmentcouncil/2020/04/29/how-blockchain-can-solve-todays-medical-supply-chain-flaws-and-improve-responses-for-future-crises/ (visited on 02/08/2023).

[49] M. Zahreddine. "Council post: How blockchain will revolutionize the pharmaceutical industry", Forbes. Section: Innovation. (), [Online]. Available: https://www.forbes.com/sites/forbestechcouncil/2018/11/14/how-blockchain-will-revolutionize-the-pharmaceutical-industry/ (visited on 02/08/2023).

[50] M. M. Aung and Y. S. Chang, "Traceability in a food supply chain: Safety and quality perspectives", *Food Control*, vol. 39, pp. 172–184, May 1, 2014, ISSN: 0956-7135. DOI: 10.1016/j.foodcont.2013.11.007. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0956713513005811 (visited on 02/08/2023).

[51] R. Badia-Melis, P. Mishra, and L. Ruiz-García, "Food traceability: New trends and recent advances. a review", *Food Control*, vol. 57, pp. 393–401, Nov. 1, 2015, ISSN: 0956-7135. DOI: 10.1016/j.foodcont.2015.05.005. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0956713515002698 (visited on 02/08/2023).

[52] E. Abad, F. Palacio, M. Nuin, *et al.*, "RFID smart tag for traceability and cold chain monitoring of foods: Demonstration in an intercontinental fresh fish logistic chain", *Journal of Food Engineering*, vol. 93, no. 4, pp. 394–399, Aug. 1, 2009, ISSN: 0260-8774. DOI: 10.1016/j.jfoodeng.2009.02.004. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0260877409000661 (visited on 02/08/2023).

[53] Z. Zou, Q. Chen, I. Uysal, and L. Zheng, "Radio frequency identification enabled wireless sensing for intelligent food logistics", *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 372, no. 2017, p. 20 130 313, Jun. 13, 2014, Publisher: Royal Society. DOI: 10.1098/rsta.2013.0313. [Online]. Available: https://royalsocietypublishing.org/doi/10.1098/rsta.2013.0313 (visited on 02/08/2023).

[54] J. Huang, X. Li, C.-C. Xing, W. Wang, K. Hua, and S. Guo, "DTD: A novel double-track approach to clone detection for RFID-enabled supply chains", *IEEE Transactions on Emerging Topics in*

Computing, vol. 5, no. 1, pp. 134–140, Jan. 2017, Conference Name: IEEE Transactions on Emerging Topics in Computing, ISSN: 2168-6750. DOI: 10.1109/TETC.2015.2389532.

[55] P. Helo, M. Suorsa, Y. Hao, and P. Anussornnitisarn, "Toward a cloud-based manufacturing execution system for distributed manufacturing", *Computers in Industry*, vol. 65, no. 4, pp. 646–656, May 1, 2014, ISSN: 0166-3615. DOI: 10.1016/j.compind.2014.01.015. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0166361514000311 (visited on 02/08/2023).

[56] N. Kshetri, "Blockchain's roles in strengthening cybersecurity and protecting privacy", *Telecommunications Policy*, Celebrating 40 Years of Telecommunications Policy – A Retrospective and Prospective View, vol. 41, no. 10, pp. 1027–1038, Nov. 1, 2017, ISSN: 0308-5961. DOI: 10.1016/j.telpol.2017.09.003. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0308596117302483 (visited on 02/08/2023).

[57] F. Tian, "A supply chain traceability system for food safety based on HACCP, blockchain & internet of things", in *2017 International Conference on Service Systems and Service Management*, ISSN: 2161-1904, Jun. 2017, pp. 1–6. DOI: 10.1109/ICSSSM.2017.7996119.

[58] G. R. Nair and S. Sebastian, "BlockChain technology centralised ledger to distributed ledger", vol. 04, no. 3,

[59] F. Tian, "An agri-food supply chain traceability system for china based on RFID & blockchain technology", in *2016 13th International Conference on Service Systems and Service Management (ICSSSM)*, ISSN: 2161-1904, Jun. 2016, pp. 1–6. DOI: 10.1109/ICSSSM.2016.7538424.

[60] "Leveraging blockchain to improve food supply chain traceability", IBM Supply Chain and Blockchain Blog. (Nov. 16, 2016), [Online]. Available: https://www.ibm.com/blogs/blockchain/2016/11/leveraging-blockchain-improve-food-supply-chain-traceability/ (visited on 10/08/2022).

[61] M. S. Christo, A. M. A., P. S. G., P. C., and R. K. M., "An efficient data security in medical report using block chain technology", in *2019 International Conference on Communication and Signal Processing (ICCSP)*, Apr. 2019, pp. 0606–0610. DOI: 10.1109/ICCSP.2019.8698058.

[62] N. Nchinda, A. Cameron, K. Retzepi, and A. Lippman, "MedRec: A network for personal information distribution", in *2019 International Conference on Computing, Networking and Communications (ICNC)*, ISSN: 2325-2626, Feb. 2019, pp. 637–641. DOI: 10.1109/ICCNC.2019.8685631.

[63] R. Kumar and R. Tripathi, "A secure and distributed framework for sharing COVID-19 patient reports using consortium blockchain and IPFS", in *2020 Sixth International Conference on Parallel, Distributed and Grid Computing (PDGC)*, ISSN: 2573-3079, Nov. 2020, pp. 231–236. DOI: 10.1109/PDGC50313.2020.9315755.

[64] M. Hanley and H. Tewari, "Managing lifetime healthcare data on the blockchain", 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)*, Oct. 2018, pp. 246–251. DOI: 10.1109/SmartWorld.2018.00077.

[65] M. T. de Oliveira, L. H. A. Reis, R. C. Carrano, *et al.*, "Towards a blockchain-based secure electronic medical record for healthcare applications", in *ICC 2019 - 2019 IEEE International Conference on Communications (ICC)*, ISSN: 1938-1883, May 2019, pp. 1–6. DOI: 10.1109/ICC.2019.8761307.

[66] 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)*, ISSN: 2325-2626, Feb. 2019, pp. 375–379. DOI: 10.1109/ICCNC.2019.8685552.

[67] 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*, vol. 5, pp. 14 757–14 767, 2017, Conference Name: IEEE Access, ISSN: 2169-3536. DOI: 10.1109/ACCESS.2017.2730843.

[68] Y. Wang, A. Zhang, P. Zhang, and H. Wang, "Cloud-assisted EHR sharing with security and privacy preservation via consortium blockchain", *IEEE Access*, vol. 7, pp. 136 704–136 719, 2019, Conference Name: IEEE Access, ISSN: 2169-3536. DOI: 10.1109/ACCESS.2019.2943153.

[69] S. Jiang, J. Cao, H. Wu, Y. Yang, M. Ma, and J. He, "BlocHIE: A BLOCkchain-based platform for healthcare information exchange", in *2018 IEEE International Conference on Smart Computing (SMARTCOMP)*, Jun. 2018, pp. 49–56. DOI: 10.1109/SMARTCOMP.2018.00073.

[70] Z. Xiao, Z. Li, Y. Liu, *et al.*, "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)*, ISSN: 1521-9097, Dec. 2018, pp. 998–1003. DOI: 10.1109/PADSW.2018.8645049.

[71] 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)*, ISSN: 2576-6813, Dec. 2018, pp. 206–212. DOI: 10.1109/GLOCOM.2018.8647221.

[72] J. Zhou, F. Tang, H. Zhu, N. Nan, and Z. Zhou, "Distributed data vending on blockchain", 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)*, Jul. 2018, pp. 1100–1107. DOI: 10.1109/Cybermatics_2018.2018.00201.

[73] S. Rouhani, L. Butterworth, A. D. Simmons, D. G. Humphery, and R. Deters, "MediChainTM: A secure decentralized medical data asset management system", 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)*, Jul. 2018, pp. 1533–1538. DOI: 10.1109/Cybermatics_2018.2018.00258.

[74] S. Alexaki, G. Alexandris, V. Katos, and N. E. 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)*, ISSN: 2378-4873, Sep. 2018, pp. 1–6. DOI: 10.1109/CAMAD.2018.8514954.

[75] C. Li, Y. Cao, Z. Hu, and M. Yoshikawa, "Blockchain-based bidirectional updates on fine-grained medical data", in *2019 IEEE 35th International Conference on Data Engineering Workshops (ICDEW)*, ISSN: 2473-3490, Apr. 2019, pp. 22–27. DOI: 10.1109/ICDEW.2019.00-40.

[76] 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", in *2018 IEEE Global Communications Conference (GLOBECOM)*, ISSN: 2576-6813, Dec. 2018, pp. 1–6. DOI: 10.1109/GLOCOM.2018.8647713.

[77] T. Mikula and R. H. Jacobsen, "Identity and access management with blockchain in electronic healthcare records", in *2018 21st Euromicro Conference on Digital System Design (DSD)*, Aug. 2018, pp. 699–706. DOI: 10.1109/DSD.2018.00008.

[78] 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)*, Mar. 2019, pp. 369–374. DOI: 10.1109/ICBDA.2019.8713188.

[79] 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)*, Oct. 2018, pp. 113–118. DOI: 10.1109/WorldS4.2018.8611621.

[80] A. R. Rajput, Q. Li, M. Taleby Ahvanooey, and I. Masood, "EACMS: Emergency access control management system for personal health record based on blockchain", *IEEE Access*, vol. 7, pp. 84 304–84 317, 2019, Conference Name: IEEE Access, ISSN: 2169-3536. DOI: 10.1109/ACCESS.2019.2917976.

[81] S. Wang, D. Zhang, and Y. Zhang, "Blockchain-based personal health records sharing scheme with data integrity verifiable", *IEEE Access*, vol. 7, pp. 102 887–102 901, 2019, Conference Name: IEEE Access, ISSN: 2169-3536. DOI: 10.1109/ACCESS.2019.2931531.

[82] X. Liu, Z. Wang, C. Jin, F. Li, and G. Li, "A blockchain-based medical data sharing and protection scheme", *IEEE Access*, vol. 7, pp. 118 943–118 953, 2019, Conference Name: IEEE Access, ISSN: 2169-3536. DOI: 10.1109/ACCESS.2019.2937685.

[83] J. Vora, A. Nayyar, S. Tanwar, *et al.*, "BHEEM: A blockchain-based framework for securing electronic health records", in *2018 IEEE Globecom Workshops (GC Wkshps)*, Dec. 2018, pp. 1–6. DOI: 10.1109/GLOCOMW.2018.8644088.

[84] 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, pp. 66 792–66 806, 2019, Conference Name: IEEE Access, ISSN: 2169-3536. DOI: 10.1109/ACCESS.2019.2917555.

[85] G. Yang and C. Li, "A design of blockchain-based architecture for the security of electronic health record (EHR) systems", in *2018 IEEE International Conference on Cloud Computing*

Technology and Science (CloudCom), ISSN: 2330-2186, Dec. 2018, pp. 261–265. DOI: 10.1109/CloudCom2018.2018.00058.

[86] 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 *2018 17th IEEE International Conference On Trust, Security And Privacy In Computing And Communications/ 12th IEEE International Conference On Big Data Science And Engineering (TrustCom/BigDataSE)*, ISSN: 2324-9013, Aug. 2018, pp. 1374–1379. DOI: 10.1109/TrustCom/BigDataSE.2018.00190.

[87] M. G. Kim, A. R. Lee, H. J. Kwon, J. W. Kim, and I. K. Kim, "Sharing medical questionnaries based on blockchain", in *2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, Dec. 2018, pp. 2767–2769. DOI: 10.1109/BIBM.2018.8621154.

[88] 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)*, Oct. 2018, pp. 257–262. DOI: 10.1109/SmartWorld.2018.00079.

[89] 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)*, ISSN: 2474-3828, Oct. 2018, pp. 829–833. DOI: 10.1109/ITME.2018.00186.

[90] U. Goel, R. Ruhl, and P. Zavarsky, "Using healthcare authority and patient blockchains to develop a tamper-proof record tracking system", in *2019 IEEE 5th Intl Conference on Big Data Security on Cloud (BigDataSecurity), IEEE Intl Conference on High Performance and Smart Computing, (HPSC) and IEEE Intl Conference on Intelligent Data and Security (IDS)*, May 2019, pp. 25–30. DOI: 10.1109/BigDataSecurity-HPSC-IDS.2019.00016.

[91] P. Li, C. Xu, H. Jin, *et al.*, "ChainSDI: A software-defined infrastructure for regulation-compliant home-based healthcare services secured by blockchains", *IEEE Systems Journal*, vol. 14, no. 2, pp. 2042–2053, Jun. 2020, Conference Name: IEEE Systems Journal, ISSN: 1937-9234. DOI: 10.1109/JSYST.2019.2937930.

[92] J. Xu, K. Xue, S. Li, *et al.*, "Healthchain: A blockchain-based privacy preserving scheme for large-scale health data", *IEEE Internet of Things Journal*, vol. 6, no. 5, pp. 8770–8781, Oct. 2019, Conference Name: IEEE Internet of Things Journal, ISSN: 2327-4662. DOI: 10.1109/JIOT.2019.2923525.

[93] V. Malamas, T. Dasaklis, P. Kotzanikolaou, M. Burmester, and S. Katsikas, "A forensics-by-design management framework for medical devices based on blockchain", in *2019 IEEE World Congress on Services (SERVICES)*, ISSN: 2642-939X, vol. 2642-939X, Jul. 2019, pp. 35–40. DOI: 10.1109/SERVICES.2019.00021.

[94] 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)*, Dec. 2018, pp. 1–6. DOI: 10.1109/GLOCOMW.2018.8644164.

[95] 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)*, Sep. 2018, pp. 1–6. DOI: 10.1109/HealthCom.2018.8531125.

[96] 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, pp. 61 876–61 885, 2018, Conference Name: IEEE Access, ISSN: 2169-3536. DOI: 10.1109/ACCESS.2018.2875242.

[97] R. Guo, H. Shi, D. Zheng, C. Jing, C. Zhuang, and Z. Wang, "Flexible and efficient blockchain-based ABE scheme with multi-authority for medical on demand in telemedicine system", *IEEE Access*, vol. 7, pp. 88 012–88 025, 2019, Conference Name: IEEE Access, ISSN: 2169-3536. DOI: 10.1109/ACCESS.2019.2925625.

[98] 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)*, Nov. 2018, pp. 1–5. DOI: 10.1109/CloudTech.2018.8713343.

[99] A. Islam and S. Y. Shin, "BHMUS: Blockchain based secure outdoor health monitoring scheme using UAV in smart city", in *2019 7th International Conference on Information and Communication Technology (ICoICT)*, Jul. 2019, pp. 1–6. DOI: 10.1109/ICoICT.2019.8835373.

[100] M. Saravanan, R. Shubha, A. M. Marks, and V. Iyer, "SMEAD: A secured mobile enabled assisting device for diabetics monitoring", in *2017 IEEE International Conference on Advanced Networks and Telecommunications Systems (ANTS)*, Dec. 2017, pp. 1–6. DOI: 10.1109/ANTS.2017.8384099.

[101] M. A. Uddin, A. Stranieri, I. Gondal, and V. Balasubramanian, "Continuous patient monitoring with a patient centric agent: A block architecture", *IEEE Access*, vol. 6, pp. 32 700–32 726, 2018, Conference Name: IEEE Access, ISSN: 2169-3536. DOI: 10.1109/ACCESS.2018.2846779.

[102] H. Kordestani, K. Barkaoui, and W. Zahran, "HapiChain: A blockchain-based framework for patient-centric telemedicine", in *2020 IEEE 8th International Conference on Serious Games and Applications for Health (SeGAH)*, ISSN: 2573-3060, Aug. 2020, pp. 1–6. DOI: 10.1109/SeGAH49190.2020.9201726.

[103] X. Liang, J. Zhao, S. Shetty, J. Liu, and D. Li, "Integrating blockchain for data sharing and collaboration in mobile healthcare applications", in *2017 IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, ISSN: 2166-9589, Oct. 2017, pp. 1–5. DOI: 10.1109/PIMRC.2017.8292361.

[104] X. Zhao, S. Wang, Y. Zhang, and Y. Wang, "Attribute-based access control scheme for data sharing on hyperledger fabric", *Journal of Information Security and Applications*, vol. 67, p. 103 182, Jun. 1, 2022, ISSN: 2214-2126. DOI: 10.1016/j.jisa.2022.103182. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2214212622000643 (visited on 03/26/2023).

[105] R. Azzi, R. K. Chamoun, and M. Sokhn, "The power of a blockchain-based supply chain", *Computers & Industrial Engineering*, vol. 135, pp. 582–592, Sep. 1, 2019, ISSN: 0360-8352. DOI: 10.1016/j.cie.2019.06.042. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0360835219303729 (visited on 09/30/2020).

[106] Q. Zhou, H. Huang, Z. Zheng, and J. Bian, "Solutions to scalability of blockchain: A survey", *IEEE Access*, vol. 8, pp. 16 440–16 455, 2020, Conference Name: IEEE Access, ISSN: 2169-3536. DOI: 10.1109/ACCESS.2020.2967218.

[107] K. Croman, C. Decker, I. Eyal, *et al.*, "On scaling decentralized blockchains", in *Financial Cryptography and Data Security*, J. Clark, S. Meiklejohn, P. Y. Ryan, D. Wallach, M. Brenner, and

K. Rohloff, Eds., ser. Lecture Notes in Computer Science, Berlin, Heidelberg: Springer, 2016, pp. 106–125, ISBN: 978-3-662-53357-4. DOI: 10.1007/978-3-662-53357-4_8.

[108] S. Thompson, "The preservation of digital signatures on the blockchain", *See Also*, no. 3, Jul. 31, 2017, Number: 3. DOI: 10.14288/sa.v0i3.188841. [Online]. Available: https://ojs.library.ubc.ca/index.php/seealso/article/view/188841 (visited on 09/30/2020).

[109] H. Editor. "What are patient rights under HIPAA?", HIPAAnswers. Section: HIPAA Questions and Answers. (Feb. 5, 2019), [Online]. Available: https://www.hipaanswers.com/patient-rights-under-hipaa/ (visited on 10/06/2020).

[110] N. Viswanadham and A. Samvedi, "Supplier selection based on supply chain ecosystem, performance and risk criteria", *International Journal of Production Research*, vol. 51, no. 21, pp. 6484–6498, Nov. 1, 2013, Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/00207543.2013.825056, ISSN: 0020-7543. DOI: 10.1080/00207543.2013.825056. [Online]. Available: https://doi.org/10.1080/00207543.2013.825056 (visited on 10/06/2022).

[111] A. Musamih, K. Salah, R. Jayaraman, *et al.*, "A blockchain-based approach for drug traceability in healthcare supply chain", *IEEE Access*, vol. 9, pp. 9728–9743, 2021, Conference Name: IEEE Access, ISSN: 2169-3536. DOI: 10.1109/ACCESS.2021.3049920.

[112] M. Guerrini, D. Beccati, Z. Shriver, *et al.*, "Oversulfated chondroitin sulfate is a contaminant in heparin associated with adverse clinical events", *Nature Biotechnology*, vol. 26, no. 6, pp. 669–675, Jun. 2008, Number: 6 Publisher: Nature Publishing Group, ISSN: 1546-1696. DOI: 10.1038/nbt1407. [Online]. Available: https://www.nature.com/articles/nbt1407 (visited on 03/07/2022).

[113] C. f. D. E. a. Research, "Information on heparin", *FDA*, Nov. 3, 2018, Publisher: FDA. [Online]. Available: https://www.fda.gov/drugs/postmarket-drug-safety-information-patients-and-providers/information-heparin (visited on 03/07/2022).

[114] "Blockchain-based traceability in agri-food supply chain management: A practical implementation | IEEE conference publication | IEEE xplore". [Online]. Available: https://ieeexplore.ieee.org/document/8373021 (visited on 02/08/2023).

[115] T. McConaghy and D. Holtzman, "Towards an ownership layer for the internet",

[116] "BlockVerify", Blockdata. (), [Online]. Available: https://www.blockdata.tech/profiles/blockverify (visited on 02/09/2023).

[117] Chronicled, Open Source, and Chronicled Open Source Team. "Chronicled white paper open registry for IOT". (), [Online]. Available: https://blockchainlab.com/pdf/whitepaper7.pdf.

[118] "OwlTing blockchain services". (), [Online]. Available: https://www.owlting.com/obs (visited on 02/09/2023).

[119] "Blockchain: The solution for supply chain transparency", Provenance. (Nov. 21, 2015), [Online]. Available: https://www.provenance.org/whitepaper (visited on 02/09/2023).

[120] "The everledger platform where supply chains meets the blockchain", Everledger. (), [Online]. Available: https://everledger.io/ (visited on 02/09/2023).

[121] "Verisart | web3 tools to sell NFTs & protect against fraud", Verisart. (), [Online]. Available: https://verisart.com/ (visited on 02/09/2023).

[122] "TrustChain", TrustChain. (), [Online]. Available: https://www.trustchainjewelry.com (visited on 02/09/2023).

[123] "Sensing system and integrity of supply chain data". (2017), [Online]. Available: https://ambrosus.com/assets/new4-3.-Sensing-System-and-Intregrity.pdf.

[124] T. Bocek, B. B. Rodrigues, T. Strasser, and B. Stiller, "Blockchains everywhere - a use-case of blockchains in the pharma supply-chain", in *2017 IFIP/IEEE Symposium on Integrated Network and Service Management (IM)*, May 2017, pp. 772–777. DOI: 10.23919/INM.2017.7987376.

[125] F. Mustafa and S. Andreescu, "Chemical and biological sensors for food-quality monitoring and smart packaging", *Foods*, vol. 7, no. 10, p. 168, Oct. 16, 2018, ISSN: 2304-8158. DOI: 10.3390/foods7100168. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6210272/ (visited on 02/09/2023).

[126] T. K. Mackey and G. Nayyar, "A review of existing and emerging digital technologies to combat the global trade in fake medicines", *Expert Opinion on Drug Safety*, vol. 16, no. 5, pp. 587–602, May 2017, ISSN: 1744-764X. DOI: 10.1080/14740338.2017.1313227.

[127] S. Al-Sarawi, M. Anbar, K. Alieyan, and M. Alzubaidi, "Internet of things (IoT) communication protocols: Review", in *2017 8th International Conference on Information Technology (ICIT)*, May 2017, pp. 685–690. DOI: 10.1109/ICITECH.2017.8079928.

[128] F. Samie, L. Bauer, and J. Henkel, "IoT technologies for embedded computing: A survey", in *2016 International Conference on Hardware/Software Codesign and System Synthesis (CODES+ISSS)*, Oct. 2016, pp. 1–10.

[129] A. Ray, V. Raj, M. Oriol, A. Monot, and S. Obermeier, "Bluetooth low energy devices security testing framework", in *2018 IEEE 11th International Conference on Software Testing, Verification and Validation (ICST)*, Apr. 2018, pp. 384–393. DOI: 10.1109/ICST.2018.00045.

[130] G. Jain and S. Dahiya, "NFC: Advantages, limits and future scope", *International Journal on Cybernetics & Informatics*, vol. 4, no. 4, pp. 1–12, Aug. 31, 2015, ISSN: 23208430, 2277548X. DOI: 10.5121/ijci.2015.4401. [Online]. Available: http://www.airccse.org/journal/ijci/papers/4415ijci01.pdf (visited on 02/09/2023).

[131] S. Krishnaswamy, "Wireless application protocol", 2001.

[132] O. Changqing, W. Jixiong, L. Zhengyan, and H. Shengye, "An enhanced security authentication protocol based on hash-lock for low-cost RFID", in *Security and Identification 2008 2nd International Conference on Anti-counterfeiting*, ISSN: 2163-5056, Aug. 2008, pp. 416–419. DOI: 10.1109/IWASID.2008.4688440.

[133] G. Madlmayr, J. Langer, C. Kantner, and J. Scharinger, "NFC devices: Security and privacy", in *2008 Third International Conference on Availability, Reliability and Security*, Mar. 2008, pp. 642–647. DOI: 10.1109/ARES.2008.105.

[134] C. Xenakis, D. Apostolopoulou, A. Panou, and I. Stavrakakis, "A qualitative risk analysis for the GPRS technology", in *2008 IEEE/IFIP International Conference on Embedded and Ubiquitous Computing*, vol. 2, Dec. 2008, pp. 61–68. DOI: 10.1109/EUC.2008.123.

[135] T. McConaghy, R. Marques, A. Muller, *et al.*, "BigchainDB: A scalable blockchain database",

[136] "Ethereum", ethereum.org. (), [Online]. Available: https://ethereum.org (visited on 02/09/2023).

[137] "Hyperledger fabric", Hyperledger. (), [Online]. Available: https://www.hyperledger.org/use/fabric (visited on 07/29/2021).

[138] LinuxFoundationX: LFS171x. "Blockchain for business - an introduction to hyperledger technologies". (), [Online]. Available: https://courses.edx.org/courses/course-v1:LinuxFoundationX+LFS171x+3T2017/course/ (visited on 12/17/2020).

[139] A. Baliga, "Understanding blockchain consensus models", 2017. [Online]. Available: https://www.semanticscholar.org/paper/Understanding-Blockchain-Consensus-Models-Baliga/da8a37b10bc1521a4d3de925d7ebc44bb606d740 (visited on 02/09/2023).

[140] E. Ramia, R. M. Zeenny, S. Hallit, P. Salameh, and on behalf of the Order of Pharmacists Scientific Committee – Medication Safety Subcommittee, "Assessment of patients' knowledge and practices regarding their medication use and risks in lebanon", *International Journal of Clinical Pharmacy*, vol. 39, no. 5, pp. 1084–1094, Oct. 1, 2017, ISSN: 2210-7711. DOI: 10.1007/s11096-017-0517-4. [Online]. Available: https://doi.org/10.1007/s11096-017-0517-4 (visited on 12/06/2020).

[141] M. Karami and A. Rahimi, "Semantic web technologies for sharing clinical information in health care systems", *Acta Informatica Medica*, vol. 27, no. 1, pp. 4–7, Mar. 2019, ISSN: 0353-8109. DOI: 10.5455/aim.2019.27.4-7. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6511266/ (visited on 06/21/2020).

[142] S. Iftikhar, F. Ahmad, and K. Fatima, "A framework based on OWL-s for healthcare information provision", in *2011 7th International Conference on Emerging Technologies*, Sep. 2011, pp. 1–6. DOI: 10.1109/ICET.2011.6048449.

[143] "Medication errors and adverse drug events", [Online]. Available: https://psnet.ahrq.gov/primer/medication-errors-and-adverse-drug-events (visited on 05/04/2022).

[144] Office of Disease Prevention and Health Promotion. "Adverse drug events | health.gov". (Feb. 5, 2020), [Online]. Available: https://health.gov/our-work/health-care-quality/adverse-drug-events (visited on 07/30/2020).

[145] R. W. Pretorius, G. Gataric, S. K. Swedlund, and J. R. Miller, "Reducing the risk of adverse drug events in older adults", *American Family Physician*, vol. 87, no. 5, pp. 331–336, Mar. 1, 2013, ISSN: 0002-838X, 1532-0650. [Online]. Available: https://www.aafp.org/afp/2013/0301/p331.html (visited on 05/04/2022).

[146] Z. Nakhla, K. Nouira, and A. Ferchichi, "Prescription adverse drug events system (PrescADE) based on ontology and internet of things", *The Computer Journal*, vol. 62, no. 6, pp. 801–805, Jun. 1, 2019, ISSN: 0010-4620. DOI: 10.1093/comjnl/bxy076. [Online]. Available: https://doi.org/10.1093/comjnl/bxy076 (visited on 05/04/2022).

[147] B. Aldughayfiq and S. Sampalli, "Patients', pharmacists', and prescribers' attitude toward using blockchain and machine learning in a proposed ePrescription system: Online survey", *JAMIA Open*, vol. 5, no. 1, ooab115, Apr. 1, 2022, ISSN: 2574-2531. DOI: 10.1093/jamiaopen/ooab115. [Online]. Available: https://doi.org/10.1093/jamiaopen/ooab115 (visited on 05/05/2022).

[148] C. Thatcher and S. Acharya, "Pharmaceutical uses of blockchain technology", in *2018 IEEE International Conference on Advanced Networks and Telecommunications Systems (ANTS)*, ISSN: 2153-1684, Dec. 2018, pp. 1–6. DOI: 10.1109/ANTS.2018.8710154.

[149] R. G. Nguewo Ngassam, R. Ologeanu-Taddei, J. Lartigau, and I. Bourdon, "A use case of blockchain in healthcare: Allergy card", in *Blockchain and Distributed Ledger Technology Use Cases: Applications and Lessons Learned*, ser. Progress in IS, H. Treiblmaier and T. Clohessy, Eds., Cham: Springer International Publishing, 2020, pp. 69–94, ISBN: 978-3-030-44337-5. DOI: 10.1007/978-3-030-44337-5_4. [Online]. Available: https://doi.org/10.1007/978-3-030-44337-5_4 (visited on 07/30/2020).

[150] I. Mitchell and S. Hara, "BMAR – blockchain for medication administration records", in *Blockchain and Clinical Trial: Securing Patient Data*, ser. Advanced Sciences and Technologies for Security Applications, H. Jahankhani, S. Kendzierskyj, A. Jamal, G. Epiphaniou, and H. Al-Khateeb, Eds., Cham: Springer International Publishing, 2019, pp. 231–248, ISBN: 978-3-030-11289-9. DOI: 10.1007/978-3-030-11289-9_10. [Online]. Available: https://doi.org/10.1007/978-3-030-11289-9_10 (visited on 07/31/2020).

[151] P. Li, S. D. Nelson, B. A. Malin, and Y. Chen, "DMMS: A decentralized blockchain ledger for the management of medication histories", *Blockchain in healthcare today*, vol. 2, 2019, ISSN: 2573-8240. DOI: 10.30953/bhty.v2.38. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7286573/ (visited on 07/31/2020).

[152] R. D. Garcia, G. A. Zutião, G. Ramachandran, and J. Ueyama, "Towards a decentralized e-prescription system using smart contracts", in *2021 IEEE 34th International Symposium on*

Computer-Based Medical Systems (CBMS), ISSN: 2372-9198, Jun. 2021, pp. 556–561. DOI: 10.1109/CBMS52027.2021.00037.

[153] A. Abatal, H. Khallouki, and M. Bahaj, "A semantic smart interconnected healthcare system using ontology and cloud computing", in *2018 4th International Conference on Optimization and Applications (ICOA)*, Apr. 2018, pp. 1–5. DOI: 10.1109/ICOA.2018.8370595.

[154] S. Subbulakshmi, A. Krishnan, and R. Sreereshmi, "Contextual aware dynamic healthcare service composition based on semantic web ontology", in *2019 2nd International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICICT)*, vol. 1, Jul. 2019, pp. 1474–1479. DOI: 10.1109/ICICICT46008.2019.8993303.

[155] T. Sigwele, Y. F. Hu, M. Ali, J. Hou, M. Susanto, and H. Fitriawan, "An intelligent edge computing based semantic gateway for healthcare systems interoperability and collaboration", in *2018 IEEE 6th International Conference on Future Internet of Things and Cloud (FiCloud)*, Aug. 2018, pp. 370–376. DOI: 10.1109/FiCloud.2018.00060.

[156] N. Karthik and V. Ananthanarayana, "A trust model for lightweight semantic annotation of sensor data in pervasive environment", in *2018 IEEE/ACIS 17th International Conference on Computer and Information Science (ICIS)*, Jun. 2018, pp. 28–33. DOI: 10.1109/ICIS.2018.8466471.

[157] S. Sabra, M. Alobaidi, K. M. Malik, and V. Sabeeh, "Performance evaluation for semantic-based risk factors extraction from clinical narratives", in *2018 IEEE 8th Annual Computing and Communication Workshop and Conference (CCWC)*, Jan. 2018, pp. 695–701. DOI: 10.1109/CCWC.2018.8301742.

[158] A. Rhayem, M. B. A. Mhiri, and F. Gargouri, "HealthIoT ontology for data semantic representation and interpretation obtained from medical connected objects", in *2017 IEEE/ACS 14th International Conference on Computer Systems and Applications (AICCSA)*, ISSN: 2161-5330, Oct. 2017, pp. 1470–1477. DOI: 10.1109/AICCSA.2017.171.

[159] H. Q. Yu, X. Zhao, Z. Deng, and F. Dong, "Semantic lifting and reasoning on the personalised activity big data repository for healthcare research", in *2017 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)*, Jun. 2017, pp. 818–823. DOI: 10.1109/iThings-GreenCom-CPSCom-SmartData.2017.125.

[160] M. A. Jarwar, S. Ali, and I. Chong, "Exploring web objects enabled data-driven microservices for e-health service provision in IoT environment", in *2018 International Conference on Information and Communication Technology Convergence (ICTC)*, ISSN: 2162-1233, Oct. 2018, pp. 112–117. DOI: 10.1109/ICTC.2018.8539684.

[161] A. Dridi, S. Sassi, and S. Faiz, "Towards a semantic medical internet of things", in *2017 IEEE/ACS 14th International Conference on Computer Systems and Applications (AICCSA)*, ISSN: 2161-5330, Oct. 2017, pp. 1421–1428. DOI: 10.1109/AICCSA.2017.194.

[162] V. Daliya and T. Ramesh, "Data interoperability enhancement of electronic health record data using a hybrid model", in *2019 International Conference on Smart Systems and Inventive Technology (ICSSIT)*, Nov. 2019, pp. 318–322. DOI: 10.1109/ICSSIT46314.2019.8987777.

[163] G.-W. Kim and D.-H. Lee, "Intelligent health diagnosis technique exploiting automatic ontology generation and web-based personal health record services", *IEEE Access*, vol. 7, pp. 9419–9444, 2019, Conference Name: IEEE Access, ISSN: 2169-3536. DOI: 10.1109/ACCESS.2019.2891710.

[164] H. Wang, X. Miao, and P. Yang, "Design and implementation of personal health record systems based on knowledge graph", in *2018 9th International Conference on Information Technology in Medicine and Education (ITME)*, ISSN: 2474-3828, Oct. 2018, pp. 133–136. DOI: 10.1109/ITME.2018.00039.

[165] S. Ali, M. G. Kibria, M. A. Jarwar, S. Kumar, and I. Chong, "Microservices model in WoO based IoT platform for depressive disorder assistance", in *2017 International Conference on Information and Communication Technology Convergence (ICTC)*, Oct. 2017, pp. 864–866. DOI: 10.1109/ICTC.2017.8190800.

[166] A. D. Alahmar and R. Benlamri, "SNOMED CT-based standardized e-clinical pathways for enabling big data analytics in healthcare", *IEEE Access*, vol. 8, pp. 92 765–92 775, 2020, Conference Name: IEEE Access, ISSN: 2169-3536. DOI: 10.1109/ACCESS.2020.2994286.

[167] H. Dhayne, R. Kilany, R. Haque, and Y. Taher, "SeDIE: A semantic-driven engine for integration of healthcare data", in *2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, Dec. 2018, pp. 617–622. DOI: 10.1109/BIBM.2018.8621243.

[168] "RDF - semantic web standards". (), [Online]. Available: https://www.w3.org/RDF/ (visited on 10/11/2020).

[169] V. D. Kolias, J. Stoitsis, S. Golemati, and K. S. Nikita, "Utilizing semantic web technologies in healthcare", in *Concepts and Trends in Healthcare Information Systems*, ser. Annals of Information Systems, D.-D. Koutsouris and A. A. Lazakidou, Eds., Cham: Springer International Publishing, 2014, pp. 9–19, ISBN: 978-3-319-06844-2. DOI: 10.1007/978-3-319-06844-2_2. [Online]. Available: https://doi.org/10.1007/978-3-319-06844-2_2 (visited on 10/11/2020).

[170] "RDF schema 1.1". (), [Online]. Available: https://www.w3.org/TR/rdf-schema/ (visited on 10/11/2020).

[171] "OWL web ontology language overview". (), [Online]. Available: https://www.w3.org/TR/2004/REC-owl-features-20040210/#property (visited on 10/11/2020).

[172] "Ontologies - w3c". (), [Online]. Available: https://www.w3.org/standards/semanticweb/ontology (visited on 10/11/2020).

[173] K. Alaoui and M. Bahaj, "Semantic oriented data modeling based on RDF, RDFS and OWL", in *Advanced Intelligent Systems for Sustainable Development (AI2SD'2019)*, M. Ezziyyani, Ed., ser. Advances in Intelligent Systems and Computing, Cham: Springer International Publishing, 2020, pp. 411–421, ISBN: 978-3-030-36674-2. DOI: 10.1007/978-3-030-36674-2_42.

[174] X. Li, J.-F. Martínez, and G. Rubio, "Towards a hybrid approach to context reasoning for underwater robots", *Applied Sciences*, vol. 7, no. 2, p. 183, Feb. 2017, Number: 2 Publisher: Multidisciplinary Digital Publishing Institute. DOI: 10.3390/app7020183. [Online]. Available: https://www.mdpi.com/2076-3417/7/2/183 (visited on 04/07/2021).

[175] Y. Liao, M. Lezoche, H. Panetto, and N. Boudjlida, "Why, where and how to use semantic annotation for systems interoperability", p. 9,

[176] "State of the art of semantic web for healthcare", *Procedia - Social and Behavioral Sciences*, vol. 195, pp. 1990–1998, Jul. 3, 2015, Publisher: Elsevier, ISSN: 1877-0428. DOI: 10.1016/j.sbspro.2015.06.213. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1877042815036927 (visited on 10/25/2020).

[177] N. Choi, I.-Y. Song, and H. Han, "A survey on ontology mapping", *ACM SIGMOD Record*, vol. 35, no. 3, pp. 34–41, Sep. 2006, ISSN: 0163-5808. DOI: 10.1145/1168092.1168097. [Online]. Available: https://dl.acm.org/doi/10.1145/1168092.1168097 (visited on 10/18/2020).

[178] World Health Organization. Action Programme on Essential Drugs and Vaccines, "How to investigate drug use in health facilities : Selected drug use indicators", World Health Organization, WHO/DAP/93.1 Unpublished, 1993, number-of-pages: 87. [Online]. Available: https://apps.who.int/iris/handle/10665/60519 (visited on 02/09/2023).

[179] M. Zand, X. B. Wu, and M. A. Morris, *Hands-On Smart Contract Development with Hyperledger Fabric V2*. " O'Reilly Media, Inc.", 2021.

[180] N. Gaur, A. O'Dowd, P. Novotny, L. Desrosiers, V. Ramakrishna, and S. A. Baset, *Blockchain with Hyperledger Fabric: Build decentralized applications using Hyperledger Fabric 2, 2nd Edition*. Packt Publishing Ltd, Nov. 27, 2020, 757 pp., Google-Books-ID: F_ILEAAAQBAJ, ISBN: 978-1-83921-617-6.

[181] A. Hajj, H. Sacre, S. Hallit, R. M. Zeenny, G. Sili, and P. Salameh, "Prescription and dispensing guidelines in lebanon: Initiative of the order of pharmacists of lebanon", *Journal of Pharmaceutical Policy and Practice*, vol. 13, no. 1, p. 70, Nov. 6, 2020, ISSN: 2052-3211. DOI: 10.1186/s40545-020-00273-9.

[182] National Clinical Guideline Centre (UK), *Drug Allergy: Diagnosis and Management of Drug Allergy in Adults, Children and Young People* (National Institute for Health and Clinical Excellence: Guidance). London: National Institute for Health and Care Excellence (UK), 2014. [Online]. Available: http://www.ncbi.nlm.nih.gov/books/NBK248066/ (visited on 12/28/2020).

[183] R. Elchamaa, R. Kilany, B. Dafflon, and Y. Ouzrout, "Semantic traffic data analysis for a local leader election algorithm (LLEA)", in *ATT@ECAI*, 2020.

[184] T. J. Lampoltshammer and S. Wiegand, "Improving the computational performance of ontology-based classification using graph databases", *Remote Sensing*, vol. 7, no. 7, pp. 9473–9491, Jul. 2015, Number: 7 Publisher: Multidisciplinary Digital Publishing Institute. DOI: 10.3390/rs70709473. [Online]. Available: https://www.mdpi.com/2072-4292/7/7/9473 (visited on 04/14/2021).

[185] R. Angles and C. Gutierrez, "Survey of graph database models", *ACM Computing Surveys*, vol. 40, no. 1, 1:1–1:39, Feb. 22, 2008, ISSN: 0360-0300. DOI: 10.1145/1322432.1322433. [Online]. Available: https://doi.org/10.1145/1322432.1322433 (visited on 04/15/2021).

[186] "DB-engines ranking of graph DBMS.", DB-Engines. (), [Online]. Available: https://db-engines.com/en/ranking/graph+dbms (visited on 10/26/2020).

[187] "DB-engines ranking of RDF stores.", DB-Engines. (), [Online]. Available: https://db-engines.com/en/ranking/rdf+store (visited on 04/16/2021).

[188] "OpenLink software: Virtuoso homepage". (), [Online]. Available: https://virtuoso.openlinksw.com/ (visited on 10/27/2020).

[189] "Introduction to the semantic web — GraphDB SE 9.7.0 documentation". (), [Online]. Available: https://graphdb.ontotext.com/documentation/standard/introduction-to-semantic-web.html#introduction-to-semantic-web-reasoning-strategies (visited on 04/18/2021).

[190] "Scalable reasoning for knowledge bases subject to changes", Ph.D. dissertation, Old Dominion University Libraries, 2014. DOI: 10.25777/SH7K-7A32. [Online]. Available: https://digitalcommons.odu.edu/computerscience_etds/65/ (visited on 04/18/2021).

[191] *6 ways pharmaceutical companies are using big data to drive innovation & value*. [Online]. Available: https://www.iqpc.com/media/1001534/35903.pdf.

[192] H. Dhayne, R. Kilany, R. Haque, and Y. Taher, "Emr2vec: Bridging the gap between patient data and clinical trial", *Computers & Industrial Engineering*, vol. 156, p. 107 236, 2021.

[193] "Medaverse – web 3.0 can help patients profit from their medical data", BeInCrypto. (Jun. 23, 2022), [Online]. Available: https://beincrypto.com/medaverse-web-3-0-can-help-patients-profit-from-their-medical-data/ (visited on 10/16/2022).

[194] G. A. Fowler, "You agreed to what? doctor check-in software harvests your health data.", *Washington Post*, Jun. 14, 2022, ISSN: 0190-8286. [Online]. Available: https://www.washingtonpost.com/technology/2022/06/13/health-privacy/ (visited on 11/06/2022).

[195] F. F. Ozair, N. Jamshed, A. Sharma, and P. Aggarwal, "Ethical issues in electronic health records: A general overview", *Perspectives in Clinical Research*, vol. 6, no. 2, pp. 73–76, 2015, ISSN: 2229-3485. DOI: 10.4103/2229-3485.153997. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4394583/ (visited on 06/15/2020).

[196] "Ransomware attack forces french hospital to transfer patients", BleepingComputer. (), [Online]. Available: https://www.bleepingcomputer.com/news/security/ransomware-attack-forces-french-hospital-to-transfer-patients/ (visited on 12/30/2022).

[197] A. H. Seh, M. Zarour, M. Alenezi, *et al.*, "Healthcare data breaches: Insights and implications", *Healthcare*, vol. 8, no. 2, p. 133, May 13, 2020, ISSN: 2227-9032. DOI: 10.3390/healthcare8020133. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7349636/ (visited on 10/29/2021).

[198] A. K. Pandey, A. I. Khan, Y. B. Abushark, *et al.*, "Key issues in healthcare data integrity: Analysis and recommendations", *IEEE Access*, vol. 8, pp. 40 612–40 628, 2020, Conference Name: IEEE Access, ISSN: 2169-3536. DOI: 10.1109/ACCESS.2020.2976687.

[199] "45 CFR part 171 – information blocking", Code of Federal Regulations. (), [Online]. Available: https://www.ecfr.gov/current/title-45/subtitle-A/subchapter-D/part-171 (visited on 12/28/2022).

[200] "Information blocking: Eight regulatory reminders for october 6th", Health IT Buzz. (Sep. 30, 2022), [Online]. Available: https://www.healthit.gov/buzz-blog/information-blocking/information-blocking-eight-regulatory-reminders-for-october-6th (visited on 12/24/2022).

[201] R. A. Tariq and P. B. Hackert, *Patient Confidentiality*. StatPearls Publishing, Sep. 26, 2022, Publication Title: StatPearls [Internet]. [Online]. Available: https://www.ncbi.nlm.nih.gov/books/NBK519540/ (visited on 11/07/2022).

[202] "HIPAA release form", HIPAA Journal. (), [Online]. Available: https://www.hipaajournal.com/hipaa-release-form/ (visited on 11/07/2022).

[203] H. Journal. "De-identification of protected health information: 2022 update", HIPAA Journal. (Mar. 18, 2022), [Online]. Available: https://www.hipaajournal.com/de-identification-protected-health-information/ (visited on 10/23/2022).

[204] F. J. de Haro-Olmo, Á. J. Varela-Vaca, and J. A. Álvarez-Bermejo, "Blockchain from the perspective of privacy and anonymisation: A systematic literature review", *Sensors*, vol. 20, no. 24, p. 7171, Jan. 2020, Number: 24 Publisher: Multidisciplinary Digital Publishing Institute, ISSN: 1424-8220. DOI: 10.3390/s20247171. [Online]. Available: https://www.mdpi.com/1424-8220/20/24/7171 (visited on 12/17/2022).

[205] A. P. Joshi, M. Han, and Y. Wang, "A survey on security and privacy issues of blockchain technology", *Mathematical Foundations of Computing*, vol. 1, no. 2, p. 121, 2018, Company: Mathematical Foundations of Computing Distributor: Mathematical Foundations of Computing Institution: Mathematical Foundations of Computing Label: Mathematical Foundations of Computing Publisher: American Institute of Mathematical Sciences. DOI: 10.3934/mfc.2018007. [Online]. Available: https://www.aimsciences.org/article/doi/10.3934/mfc.2018007 (visited on 11/22/2021).

[206] Q. Feng, D. He, S. Zeadally, M. K. Khan, and N. Kumar, "A survey on privacy protection in blockchain system", *Journal of Network and Computer Applications*, vol. 126, pp. 45–58, Jan. 15, 2019, ISSN: 1084-8045. DOI: 10.1016/j.jnca.2018.10.020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1084804518303485 (visited on 11/22/2021).

[207] J. Bernal Bernabe, J. L. Canovas, J. L. Hernandez-Ramos, R. Torres Moreno, and A. Skarmeta, "Privacy-preserving solutions for blockchain: Review and challenges", *IEEE Access*, vol. 7, pp. 164 908–164 940, 2019, Conference Name: IEEE Access, ISSN: 2169-3536. DOI: 10.1109/ACCESS.2019.2950872.

[208] E. Androulaki, G. O. Karame, M. Roeschlin, T. Scherer, and S. Capkun, "Evaluating user privacy in bitcoin", in *Financial Cryptography and Data Security*, A.-R. Sadeghi, Ed., ser. Lecture Notes in Computer Science, Berlin, Heidelberg: Springer, 2013, pp. 34–51, ISBN: 978-3-642-39884-1. DOI: 10.1007/978-3-642-39884-1_4.

[209] H. Zhao, P. Bai, Y. Peng, and R. Xu, "Efficient key management scheme for health blockchain", *CAAI Transactions on Intelligence Technology*, vol. 3, no. 2, pp. 114–118, 2018, Conference Name: CAAI Transactions on Intelligence Technology, ISSN: 2468-2322. DOI: 10.1049/trit.2018.0014.

[210] M. Landers. "'it's evil' ransomware attack on hospital system in savannah is part of a growing trend", Savannah Morning News. (), [Online]. Available: https://www.savannahnow.com/story/news/2021/06/25/cyberattack-savannah-hospital-system-part-growing-trend/5336312001/ (visited on 10/31/2021).

[211] R. Sivan and Z. A. Zukarnain, "Security and privacy in cloud-based e-health system", *Symmetry*, vol. 13, no. 5, p. 742, May 2021, Number: 5 Publisher: Multidisciplinary Digital Publishing

Institute, ISSN: 2073-8994. DOI: 10.3390/sym13050742. [Online]. Available: https://www.mdpi.com/2073-8994/13/5/742 (visited on 09/28/2022).

[212] M. Chen, T. Malook, A. U. Rehman, *et al.*, "Blockchain-enabled healthcare system for detection of diabetes", *Journal of Information Security and Applications*, vol. 58, p. 102 771, May 1, 2021, ISSN: 2214-2126. DOI: 10.1016/j.jisa.2021.102771. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S221421262100020X (visited on 09/28/2022).

[213] Y. Lin and C. Zhang, "A method for protecting private data in IPFS", in *2021 IEEE 24th International Conference on Computer Supported Cooperative Work in Design (CSCWD)*, May 2021, pp. 404–409. DOI: 10.1109/CSCWD49262.2021.9437830.

[214] R. Azzi, R. Kilany Chamoun, A. Serhrouchni, and M. Sokhn, "A healthcare delivery system powered by semantic data description and blockchain", in *Future of Information and Communication Conference*, Springer, 2023, pp. 224–242.

[215] F. N. d. S. Vanin, L. M. Policarpo, R. d. R. Righi, *et al.*, "A blockchain-based end-to-end data protection model for personal health records sharing: A fully homomorphic encryption approach", *Sensors*, vol. 23, no. 1, p. 14, Jan. 2023, Number: 1 Publisher: Multidisciplinary Digital Publishing Institute, ISSN: 1424-8220. DOI: 10.3390/s23010014. [Online]. Available: https://www.mdpi.com/1424-8220/23/1/14 (visited on 04/17/2023).

[216] B. Alamri, I. T. Javed, and T. Margaria, "Preserving patients' privacy in medical IoT using blockchain", in *Edge Computing – EDGE 2020*, A. Katangur, S.-C. Lin, J. Wei, S. Yang, and L.-J. Zhang, Eds., ser. Lecture Notes in Computer Science, Cham: Springer International Publishing, 2020, pp. 103–110, ISBN: 978-3-030-59824-2. DOI: 10.1007/978-3-030-59824-2_9.

[217] S. L. Garfinkel, "De-identification of personal information", *NIST*, Oct. 22, 2015, Last Modified: 2020-01-27T16:24-05:00 Publisher: Simson L. Garfinkel. [Online]. Available: https://www.nist.gov/publications/de-identification-personal-information (visited on 04/22/2023).

[218] "Identity mixer - IBM". (Jul. 25, 2016), [Online]. Available: https://researcher.watson.ibm.com/researcher/view_group.php?id=8254 (visited on 04/24/2023).

[219] C. Stamatellis, P. Papadopoulos, N. Pitropakis, S. Katsikas, and W. J. Buchanan, "A privacy-preserving healthcare framework using hyperledger fabric", *Sensors*, vol. 20, no. 22, p. 6587, Jan. 2020, Number: 22 Publisher: Multidisciplinary Digital Publishing Institute, ISSN: 1424-8220. DOI: 10.3390/s20226587. [Online]. Available: https://www.mdpi.com/1424-8220/20/22/6587 (visited on 04/17/2023).

[220] "MSP implementation with identity mixer — hyperledger-fabricdocs main documentation". (), [Online]. Available: https://hyperledger-fabric.readthedocs.io/en/release-2.5/idemix.html (visited on 04/24/2023).

[221] O. Vovk, G. Piho, and P. Ross, "Anonymization methods of structured health care data: A literature review", in *Model and Data Engineering*, C. Attiogbé and S. Ben Yahia, Eds., ser. Lecture Notes in Computer Science, Cham: Springer International Publishing, 2021, pp. 175–189, ISBN: 978-3-030-78428-7. DOI: 10.1007/978-3-030-78428-7_14.

[222] Y. Chen, L. Meng, H. Zhou, and G. Xue, "A blockchain-based medical data sharing mechanism with attribute-based access control and privacy protection", *Wireless Communications and Mobile Computing*, vol. 2021, e6685762, Jul. 1, 2021, Publisher: Hindawi, ISSN: 1530-8669. DOI: 10.1155/2021/6685762. [Online]. Available: https://www.hindawi.com/journals/wcmc/2021/6685762/ (visited on 04/17/2023).

[223] B. Grover and D. K. Kushwaha, "Authorization and privacy preservation in cloud-based distributed ehr system using blockchain technology and anonymous digital ring signature", *Health Services and Outcomes Research Methodology*, Jun. 27, 2022, ISSN: 1572-9400. DOI: 10.1007/s10742-022-00281-z. [Online]. Available: https://doi.org/10.1007/s10742-022-00281-z (visited on 04/17/2023).

[224] D. Lee and M. Song, "MEXchange: A privacy-preserving blockchain-based framework for health information exchange using ring signature and stealth address", *IEEE Access*, vol. 9, pp. 158 122–158 139, 2021, Conference Name: IEEE Access, ISSN: 2169-3536. DOI: 10.1109/ACCESS.2021.3130552.

[225] "Art. 4 GDPR – definitions", General Data Protection Regulation (GDPR). (), [Online]. Available: https://gdpr-info.eu/art-4-gdpr/ (visited on 04/24/2023).

[226] *Deploying pseudonymisation techniques*, in collab. with F. Guasconi, P. Angelidis, and P. Drogkaris, Mar. 24, 2022. [Online]. Available: https://www.enisa.europa.eu/publications#c3=2013&c3=2023&c3=false&c5=publicationDate&reversed=on&b_start=0&c4=DEPLOYING+PSEUDONYMISATION+TECHNIQUES.

[227] "Qu'est-ce ce qu'une donnée de santé ? | CNIL". (), [Online]. Available: https://www.cnil.fr/fr/quest-ce-ce-quune-donnee-de-sante (visited on 05/09/2023).

[228] N. B. Sayles, A. H. I. M. Association, and L. L. Gordon, *Health Information Management Technology: An Applied Approach*, 5th ed. edition. Chicago, Illinois: American Helath Information Management Association, Jul. 1, 2016, 686 pp., ISBN: 978-1-58426-517-7.

[229] "Health insurance portability and accountability act of 1996 (HIPAA) | CDC". (Jun. 28, 2022), [Online]. Available: https://www.cdc.gov/phlp/publications/topic/hipaa.html (visited on 09/22/2022).

[230] EUR-Lex. "Regulation (eu) 2016/679 of the european parliament and of the council of 27 april 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing directive 95/46/ec (general data protection regulation)". (Apr. 27, 2016), [Online]. Available: https://eur-lex.europa.eu/eli/reg/2016/679/oj/eng (visited on 09/23/2022).

[231] *Data classification*, 2019. [Online]. Available: https://www.oas.org/en/sms/cicte/docs/ENG-Data-Classidication.pdf.

[232] S. Harris and F. Maymi, "CISSP", in 8th ed., Mc Graw Hill Education, Oct. 2018, ISBN: 978-1-260-14264-8.

[233] C. Ferris. "Does hyperledger fabric perform at scale?", IBM Blog. (Apr. 2, 2019), [Online]. Available: https://www.ibm.com/blog/does-hyperledger-fabric-perform-at-scale/ (visited on 06/13/2023).

[234] "IPFS powers the distributed web". (), [Online]. Available: https://ipfs.tech/ (visited on 11/10/2022).

[235] J. Benet, *IPFS - content addressed, versioned, p2p file system*, Jul. 14, 2014. DOI: 10.48550/arXiv.1407.3561. arXiv: 1407.3561[cs]. [Online]. Available: http://arxiv.org/abs/1407.3561 (visited on 05/28/2023).

[236] O. o. t. Commissioner. "The drug development process", FDA. Publisher: FDA. (Feb. 20, 2020), [Online]. Available: https://www.fda.gov/patients/learn-about-drug-and-device-approvals/drug-development-process (visited on 07/15/2023).

[237] "Celebrating 20 years of ClinicalTrials.gov and looking to the future – NIH extramural nexus". (Jan. 7, 2020), [Online]. Available: https://nexus.od.nih.gov/all/2020/01/07/celebrating-20-years-of-clinicaltrials-gov-and-looking-to-the-future/ (visited on 07/15/2023).

[238] *Patient Recruitment in Clinical Trials: Steps to Develop a Successful Enrollment Strategy*. Oregon Health & Science University. [Online]. Available: https://www.ohsu.edu/sites/default/files/2019-12/Forte_Patient_Recruitment_ebook_2017.pdf.

[239] H. Weber. "How to recruit patients for clinical trials - updated 2023 | patient recruitment tips", MEDICO DIGITAL. (Apr. 24, 2023), [Online]. Available: https://www.medicodigital.co.uk/how-to-recruit-patients-for-clinical-trials/ (visited on 07/16/2023).

[240] S. Rahman, M. A. A. Majumder, S. F. Shaban, *et al.*, "Physician participation in clinical research and trials: Issues and approaches", *Advances in Medical Education and Practice*, vol. 2, pp. 85–93, Mar. 7, 2011, ISSN: 1179-7258. DOI: 10.2147/AMEP.S14103. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3661249/ (visited on 07/16/2023).

[241] D. T. Chen, F. G. Miller, and D. L. Rosenstein, "Clinical research and the physician–patient relationship", *Annals of Internal Medicine*, vol. 138, no. 8, pp. 669–672, Apr. 15, 2003, Publisher: American College of Physicians, ISSN: 0003-4819. DOI: 10.7326/0003-4819-138-8-200304150-00015. [Online]. Available: https://www.acpjournals.org/doi/full/10.7326/0003-4819-138-8-200304150-00015 (visited on 07/16/2023).

[242] V. Totten, E. L. Simon, M. Jalili, and H. R. Sawe, "Acquiring data in medical research: A research primer for low- and middle-income countries", *African Journal of Emergency Medicine*, vol. 10, S135–S139, Jan. 1, 2020, ISSN: 2211-419X. DOI: 10.1016/j.afjem.2020.09.009. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2211419X20301051 (visited on 07/16/2023).

[243] A. Team. "The AVISPA project". (), [Online]. Available: https://www.avispa-project.org/ (visited on 09/05/2023).

[244] T. Team *et al.*, "Avispa v1. 1 user manual", *Information society technologies programme (June 2006), http://avispa-project. org*, vol. 62, p. 112, 2006.

[245] D. Plasto, "Automated analysis of industrial scale security protocols.", Ph.D. dissertation, Bond University, 2004. [Online]. Available: `https://research.bond.edu.au/en/studentTheses/automated-analysis-of-industrial-scale-security-protocols` (visited on 09/04/2023).

ECOLE
DOCTORALE

Titre : La Blockchain au Service de la Santé: vers un Écosystème Centré sur le Patient

Mots clés : Blockchain, Traçabilité, Web Sémantique, Donnée de Santé, Sécurité des Données, Données Privées, Gouvernance des Données

Résumé : L'écosystème de soins de santé évolue constamment, sous l'influence des avancées technologiques qui l'ont orienté vers des approches centrées sur le patient. Toutefois, cette transformation est associée à de nombreux défis en raison de la complexité inhérente et de la fragmentation du système de santé. Cette thèse présente une problématique à trois niveaux. Chaque niveau traite un défi de l'écosystème de soins de santé ayant une répercussion sur la santé des patients ou portant atteinte à leurs vies privées. Le premier défi abordé est celui des médicaments contrefaits ou falsifiés qui représentent une menace pour la santé publique. Le deuxième défi concerne la fragmentation des données de santé qui entrave la coordination des soins et nuit à l'efficacité clinique. Le troisième défi s'attaque à la confidentialité des données relatives aux patients, impliquant aussi la protection de leurs vies privées. La blockchain apparait comme une technologie prometteuse, capable de relever ces différents défis. Introduite dans l'écosystème de santé, la blockchain a le potentiel de renforcer la transparence, l'authentification, la sécurité et la fiabilité. Néanmoins, cette technologie s'accompagne également de son lot de défis. Cette thèse évalue les risques et opportunités liés à l'adoption de la blockchain dans l'écosystème de soins de santé. Nous commençons par une étude approfondie sur le rôle de la blockchain à améliorer la gestion de la chaîne d'approvisionnement et de la chaîne de prestation de soins de santé. Pour compléter cette approche théorique, nous intégrons des applications concrètes du monde réel afin d'élaborer les exigences nécessaires à établir une chaine d'approvisionnement basée sur la blockchain. Notre troisième contribution, présente une approche axée sur le patient, où nous combinons la technologie blockchain et les technologies du Web sémantique pour aider les patients à gérer leurs données de santé. Notre quatrième contribution s'inscrit dans le cadre de la gouvernance des données. Nous développons un Framework basé sur la blockchain pour améliorer la sécurité des données et qui par la suite pourra être adopter dans divers domaines.

Title : Blockchain Adoption in Healthcare: Toward a Patient Centric Ecosystem

Keywords : Blockchain, Traceability, Semantic Web, Health Data, Data Security, Data Privacy, Data Governance

Abstract : The healthcare sector evolves constantly, driven by technological advancement and innovative solutions. These technologies have shifted the healthcare ecosystem to be more patient-centered, focusing on meeting the patient's needs rather than the needs of the individual organizations within it. However, this transformative shift experienced by the healthcare industry is associated with multiple challenges due to the inherent complexity and fragmentation of the healthcare ecosystem. This dissertation addresses three healthcare ecosystem challenges that significantly impact patients. The first challenge addressed is the problem of counterfeit or falsified drugs that represent a threat to public health. The second challenge addressed is the problem of healthcare data fragmentation that thwarts care coordination and impacts clinical efficiency. The third challenge addressed is the confidentiality and privacy of healthcare data that, if compromised, shatter the trust relationship between patients and healthcare stakeholders.
Blockchain has emerged as a promising solution to address these critical challenges. It was introduced into the healthcare ecosystem with the promise of enforcing transparency, authentication, security, and trustworthiness. Through comprehensive analysis and case studies, this dissertation assesses the opportunities and addresses the challenges of adopting the blockchain in the healthcare industry. We start with a thorough review of the state of the art covering the blockchain's role in improving supply chain management and enhancing the healthcare delivery chain. Second, we combine theoretical and real-world application studies to develop a guideline that outlines the requirements for building a blockchain-based supply chain. Third, we propose a patient-centric framework that combines blockchain technology with Semantic technologies to help patients manage their health data. Our fourth contribution presents a novel approach to data governance by developing a blockchain-based framework that improves data security and empowers patients to participate actively in their healthcare decisions. In this final contribution, we widen the scope of the proposed framework to include a roadmap for its adoption across diverse domains.

Institut Polytechnique de Paris
91120 Palaiseau, France