

A Security and Privacy Compliant Data Sharing Solution For Healthcare Data Ecosystems

by

Raza Nowrozy

A thesis submitted in total fulfillment for the
degree of Doctor of Philosophy

Victoria University, Australia
Institute of Sustainable Industries & Liveable Cities

April 2024

Abstract

In the evolving landscape of healthcare, the *complexity* and *digitization* of medical data necessitate robust **Electronic Health Records (EHR)** systems, capable of mitigating increasing cybersecurity threats without undermining patient care. This thesis introduces a **CEMPS framework (Centralised EHR Model for Preserving Privacy and Security)**, developed in response to vulnerabilities in EHR systems. CEMPS aims to safeguard sensitive health information across healthcare spectrum, including medical care, pharmaceuticals, and health insurance.

Adopting a holistic approach, the study explores *privacy* and *security* standards, aligning health information classification with regulations such as the Australian Privacy Acts, The Australian Privacy Principles (APPs), Health Insurance Portability and Accountability Act (HIPAA), and the EU General Data Protection Regulation (GDPR). The CEMPS integrates strict security policies and advanced privacy techniques to facilitate secure health data exchange among key stakeholders like doctors, nurses, and researchers, crucial for optimizing health outcomes and efficiency.

The thesis further explores the CEMPS framework through a theoretical lens, focusing on its design principles and the mechanisms it employs to improve privacy and security within EHR systems. This theoretical examination underscores the framework's capacity to ensure robust protection of sensitive health information, leveraging rational arguments to advocate for its efficacy. By emphasizing the strategic alignment of CEMPS with prevailing privacy standards and security protocols, this analysis illustrates how the framework can significantly elevate the management, *privacy*, *security* and *confidentiality* of EHR systems, offering a more controlled environment for health data.

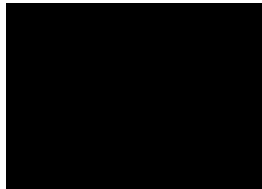
Ultimately, this thesis advocates for the industry-wide adoption of CEMPS, promoting a *secure*, *efficient*, and *privacy-compliant* healthcare environment. This research represents a significant step towards a healthcare landscape where EHR systems are both protectors of patient data and facilitators of improved healthcare delivery.

Declaration of Authorship

I, Raza Nowrozy, declare that the PhD thesis entitled **A Security and Privacy Compliant Data Sharing Solution for Healthcare Data Ecosystems** is no more than 80,000 words in length including quotes and exclusive of tables, figures, appendices, bibliography, references and footnotes. This thesis contains no material that has been submitted previously, in whole or in part, for the award of any other academic degree or diploma. Except where otherwise indicated, this thesis is my own work.

I have conducted my research in alignment with the Australian Code for the Responsible Conduct of Research and Victoria University's Higher Degree by Research Policy and Procedures.

Signed:



Date: **07/02/2024**

Acknowledgements

First of all, my deepest and most profound gratitude is extended to my primary supervisor, A/Professor Khandakar Ahmed, whose guidance and mentorship have been the cornerstone of my academic journey. A/Professor Ahmed's expertise, patience, and constant support have profoundly shaped my research, providing me with the skills and confidence to navigate through the complexities of my study. His insightful feedback, encouragement, and wisdom have been invaluable and for this I am eternally grateful.

Equally, I am immensely grateful to my co-supervisor, Professor Hua Wang, whose expertise and keen insights have greatly contributed to the depth and quality of my research. Professor Wang's dedicated mentorship and astute guidance have been instrumental in refining my academic skills and enhancing my understanding of the field. His ability to challenge and inspire has been a vital part of my academic development.

I must also express my heartfelt thanks to Dr. Tim McIntosh, whose inspirational guidance and support with my publications have been crucial. Dr. McIntosh's mentorship has motivated me to pursue excellence in my academic work, providing me with invaluable lessons that extend beyond academia.

Additionally, my sincere appreciation goes to Professor Paul Watters for his valuable insights and guidance throughout this journey. His expertise has been a key factor in broadening my perspective and deepening my understanding of the field.

My special thanks are extended to my wife, Mina Zargarzadah, for her endless support and encouragement, and to my children, Nikki and Negin Nowrozy, whose love and understanding have lightened the burden of my academic pursuits. Their constant support has profoundly influenced my journey in both academic and personal matters.

Lastly, I am grateful to Lubna Meer for her mentorship and advice, which have been instrumental in the successful completion of this thesis.

The contributions of each of these individuals have been integral to my academic journey and I acknowledge their role in the completion of this thesis with deep gratitude.

Details of Included Papers



DETAILS OF INCLUDED PAPERS: THESIS WITH PUBLICATION

Please list details of each scholarly publication and/or manuscript included in the thesis submission. Copies of published scholarly publications and/or manuscripts submitted and/or final draft manuscripts should also be included in the thesis submission.

This table must be incorporated in the thesis before the Table of Contents.

| Chapter No. | Publication Title | Publication Status | Publication Details |
|-------------|---|---|--|
| | | <ul style="list-style-type: none"> Published Accepted for publication In revised and resubmit stage Under review Manuscript ready for submission | <ul style="list-style-type: none"> Citation, if published Title, Journal, Date of acceptance letter and Corresponding editor's email address Title, Journal, Date of submission |
| 4 PART B: | ENHANCING HEALTH INFORMATION SYSTEM SECURITY: AN ONTOLOGY MODEL APPROACH | Published | Citation: 1 Title: ENHANCING HEALTH INFORMATION SYSTEM SECURITY: AN ONTOLOGY MODEL APPROACH Journal: Springer |
| 5 | TOWARDS A UNIVERSAL PRIVACY MODEL FOR ELECTRONIC HEALTH RECORD SYSTEMS: AN ONTOLOGY AND MACHINE LEARNING APPROACH | Published | Citation: 7 Title: TOWARDS A UNIVERSAL PRIVACY MODEL FOR ELECTRONIC HEALTH RECORD SYSTEMS: AN ONTOLOGY AND MACHINE LEARNING APPROACH |
| | | | |
| | | | |
| | | | |
| | | | |

Declaration by [candidate name]:

Signature:

Date:

Updated: September 2020



PRIVACY STATEMENT Victoria University (VU) values your privacy and is committed to handling your personal information in accordance with the Privacy and Data Protection Act 2014 (Vic) and other applicable privacy legislation. The personal information collected on this form will be used primarily for the purposes of assessing and processing this application. VU may also use and disclose your personal information to verify the information provided by you, to comply with government and other reporting requirements and/or to carry out associated activities connected with this application. Your personal information may also be disclosed to Commonwealth and State agencies such as the Department of Education and Training and the Department of Home Affairs in accordance with VU's obligations under the Education Services for Overseas Students Act 2000 (Cth) (ESOS Act), the National Code of Practice for Providers of Education and Training to Overseas Students 2018 (National Code) and other applicable legislation. Your personal information will not otherwise be used or disclosed without your consent, unless permitted by law. By completing and submitting this application, you agree to VU collecting, using and disclosing your personal information as described above and in accordance with VU's Privacy Policy and Student Information Privacy Collection Statement (which provides further detail about the types of personal information VU may collect from you and how it is managed) available on the Privacy page on our website vu.edu.au/privacy. You have a right to access your personal information held by VU. If you have any questions regarding privacy, please refer to the Privacy page on our website, our frequently asked questions at ASKVU or phone us on 9919 6100 or 1300 VIC UNI (or 1300 842 864).

PRIVACY INFORMATION: We collect and protect your personal information in accordance with our Privacy Policy vu.edu.au/privacy.

Victoria University CRICOS Provider No. 00124K (Melbourne) and CRICOS Provider No. 02475D (Sydney). RTO Code: 3113. ABN: 83 776 954 731

List of Publications

This dissertation brings together my own research efforts, much of which the academic community has already seen or is currently reviewing. I promise that these works are original, and I have made sure to get the right permissions from publishers to use this material in my dissertation. In this work, I discuss data sets, ideas, opinions, and diagrams that have been published before. I have carefully credited these sources to maintain academic honesty.

In this dissertation, I acknowledge the important role these previous publications have played. Their contribution has helped shape and refine the ideas and methods in my research. Recognizing this, I highlight their significant influence on the academic journey of my dissertation.

Acknowledgment of Publications

Published:

- Nowrozy, R., et al. (2023). *Towards a Universal Privacy Model for Electronic Health Record Systems*. Informatics, 10(3), 60. <https://doi.org/10.3390/informatics10030060>. Incorporated as Chapter 5.
- Nowrozy, R., et al. (2023). *Enhancing Health Information Systems Security*. In Health Information Science, HIS 2023, LNCS 14305, Springer. https://doi.org/10.1007/978-981-15-7108-4_8. Incorporated as Chapter 4 Part B.
- Nowrozy, R., et al. (2023). *Privacy Preservation of Electronic Health Records*. ACM Computing Surveys. <https://dl.acm.org/doi/abs/10.1145/3653297> Incorporated as Chapter 2.

Under Peer Review:

- *GPT, Ontology, and CAABAC: A Tripartite Personalized Access Control Model Anchored by Compliance, Context and Attribute*. Submitted to *PLOS ONE* (<https://journals.plos.org/plosone/>) and is currently under Peer Review. Incorporated as Chapter 4 Part A.

Contents

| | |
|--|--------------|
| Abstract | i |
| Declaration of Authorship | ii |
| Acknowledgements | iii |
| Details of Included Papers | iv |
| List of Publications | vi |
| List of Tables | xv |
| List of Figures | xvi |
| List of Abbreviations | xviii |
| 1 INTRODUCTION | 1 |
| 1.1 Comparative Advantages of EHRs | 3 |
| 1.1.1 Electronic Health Records Benefits | 4 |
| 1.2 Synergistic Integration of Technologies for EHR Enhancement | 5 |
| 1.2.1 EHR Drawbacks | 6 |
| 1.3 Confidentiality, Integrity, and Availability (CIA) in EHR | 7 |
| Confidentiality | 7 |
| Integrity | 8 |
| Availability | 9 |
| | 9 |
| 1.3.1 Prevalence and Impact of Cyber Attacks on Healthcare Systems | 9 |
| 1.3.2 Comparative Analysis of Global Healthcare Data Policies | 10 |
| HIPAA | 10 |
| GDPR | 10 |
| Comparative Challenges | 10 |
| | 11 |
| 1.3.3 Academic Perspectives in EHR | 11 |
| 1.4 Privacy Preservation of EHR | 12 |
| 1.5 Centralised vs Decentralized EHR System | 13 |
| 1.5.1 Rationale for Selecting a Centralized EHR System | 14 |
| 1.5.2 Justification for Centralized EHR Systems | 16 |
| Enhanced Interoperability | 16 |
| Improved Data Integrity and Quality | 16 |

| | | |
|----------|---|-----------|
| | Streamlined Governance and Compliance | 16 |
| | Real-World Efficacy | 16 |
| | | 17 |
| 1.6 | Research Motivation | 17 |
| 1.7 | Problem Statement | 18 |
| 1.8 | Research Questions | 19 |
| 1.9 | Research Aim | 20 |
| 1.10 | Research Objectives | 22 |
| | 1.10.1 Comparative Analysis and Evaluation Chapter 5 | 23 |
| 1.11 | Contribution of the Thesis | 23 |
| 1.12 | Novelty and Superiority of the Research | 24 |
| | 1.12.1 Systematic Survey of Privacy Preservation Methods Chapter 2 | 24 |
| | 1.12.2 CEMPS Framework: A Paradigm Shift in EHR Privacy and Security Chapters 3 and 4 | 25 |
| 1.13 | Thesis Outline | 25 |
| 2 | PRIVACY PRESERVATION OF ELECTRONIC HEALTH RECORDS IN THE MODERN ERA: A SYSTEMATIC SURVEY | 27 |
| 2.1 | Introduction | 27 |
| 2.2 | Background | 28 |
| | 2.2.1 EHR | 29 |
| | 2.2.2 EHR Privacy | 30 |
| | 2.2.3 EHR Confidentiality | 31 |
| | 2.2.4 EHR Security | 31 |
| | 2.2.5 The Distinction Between EHR Privacy, Confidentiality and Security | 32 |
| 2.3 | The Survey Plan and Conduct | 34 |
| | 2.3.1 Planning the Survey | 34 |
| | 2.3.1.1 Constructing the Survey Questions | 34 |
| | 2.3.1.2 The Search Keywords | 35 |
| | 2.3.1.3 Data Collection | 36 |
| | 2.3.1.4 Approach to Gathering Data | 36 |
| | 2.3.2 Conducting Survey | 37 |
| | 2.3.2.1 Executing Survey | 37 |
| | 2.3.2.2 Data Synthesis | 37 |
| 2.4 | Survey Results | 38 |
| | 2.4.1 What EHR data sharing methods are currently available? | 38 |
| | 2.4.1.1 Cloud-based Sharing | 38 |
| | 2.4.1.2 Attribute-based Access Control (ABAC) | 39 |
| | 2.4.1.3 Role-based Access Control (RBAC) | 39 |
| | 2.4.1.4 Encryption-based Sharing | 39 |
| | 2.4.1.5 Blockchain-based Sharing | 40 |
| | 2.4.1.6 HIPPA/Privacy Act-based Sharing | 41 |
| | 2.4.1.7 Current <i>Australian MyHR (My Health Record)</i> Sharing | 41 |
| | 2.4.1.8 Model-based Sharing | 41 |
| | 2.4.2 What role does privacy play when sharing EHR with different stakeholders? | 42 |

| | | |
|----------|--|-----------|
| 2.4.2.1 | Differentiation of Health Record Sets | 43 |
| 2.4.2.2 | Case Studies in EHR Data Security | 44 |
| 2.4.3 | What are the main strengths and weaknesses of EHR? | 46 |
| 2.4.3.1 | Strengths of EHR | 47 |
| 2.4.3.2 | Weaknesses of EHR | 47 |
| 2.4.4 | What is the difference between EHR Privacy, Confidentiality and Security? | 48 |
| 2.4.5 | What different technologies are available to preserve the privacy of EHRs? | 52 |
| 2.4.5.1 | Access Control Techniques | 53 |
| 2.4.5.2 | Blockchain Techniques | 56 |
| 2.4.5.3 | Cloud-based Techniques | 61 |
| 2.4.5.4 | Cryptography Techniques | 61 |
| 2.4.6 | Enhancing EHR Security with Multi-Factor Authentication | 62 |
| 2.5 | SURVEY FINDINGS | 63 |
| 2.5.1 | Privacy, Confidentiality and Security: the Difference | 63 |
| 2.5.2 | EHR Privacy Concerns | 64 |
| 2.5.3 | Discussion | 67 |
| 2.5.4 | Limitations | 68 |
| 2.6 | Conclusion and Future Research Directions | 71 |
| 3 | A SECURITY AND PRIVACY COMPLAINT DATA SHARING SOLUTION FOR HEALTHCARE DATA ECOSYSTEMS: CEMPS (CENTRALIZED EHR MODEL TO PRESERVE PRIVACY AND SECURITY) | 72 |
| 3.1 | Introduction | 72 |
| 3.2 | Literature Review | 75 |
| 3.3 | EHR Generic Architecture | 76 |
| 3.3.1 | EHR System Architecture | 76 |
| 3.3.2 | EHR System Implementation | 76 |
| 3.3.3 | Methodological Assessment and Framework Validation | 78 |
| 3.4 | Proposed CEMPS Architecture | 79 |
| 3.4.1 | Data Storage Layer | 82 |
| 3.4.1.1 | Data Storage Layer with FL | 82 |
| 3.4.1.2 | Encryption and Secure Data Transmission Protocols | 83 |
| 3.4.1.3 | Improved Privacy and Security in Centralized EHR Systems by FL | 85 |
| 3.4.1.4 | Analysing Data Storage Layer with FL and Cryptography | 86 |
| 3.4.2 | Data Sharing and Access Layer | 87 |
| 3.4.2.1 | Data Sharing and Access Layer with DP | 88 |
| 3.4.2.2 | Analysis of DP at Data Sharing and Access Layer | 89 |
| 3.4.2.3 | Mathematical Proof of Enhanced Privacy and Security | 90 |
| 3.4.3 | Centralized vs. Decentralized EHR Systems | 91 |
| 3.5 | Proposed CEMPS Methodological Framework | 91 |
| 3.5.1 | Identification Layer | 91 |
| 3.5.2 | Modeling Layer | 93 |
| 3.5.3 | Implementation Layer | 93 |

| | | |
|--|---|------------|
| 3.5.4 | Policy, Regulation, and Ethical Considerations in Centralized EHR Systems | 94 |
| 3.5.5 | Efficacy of Centralized and Decentralized EHR Models in Healthcare | 95 |
| 3.6 | Evaluation Layer | 96 |
| 3.6.1 | Comparison with Existing Privacy Preservation Techniques | 98 |
| 3.6.1.1 | Traditional Privacy Preservation Techniques | 98 |
| 3.6.1.2 | Advanced Privacy-Enhancing Technologies (PETs) | 98 |
| 3.6.1.3 | CEMPS Framework Advantages | 99 |
| 3.6.2 | Critical Evaluation of CEMPS | 99 |
| 3.7 | Limitations | 100 |
| 3.8 | Future Work | 101 |
| 3.9 | Conclusion | 102 |
| 4 | INTEGRATING ADVANCED TECHNOLOGIES AND ONTOLOGY MODELS FOR ENHANCED SECURITY IN ELECTRONIC HEALTH RECORDS | 104 |
| | Introduction | 104 |
| | | |
| PART A: GPT, ONTOLOGY, AND CAABAC: A TRIPARTITE PERSONALIZED ACCESS CONTROL MODEL ANCHORED BY COMPLIANCE, CONTEXT AND ATTRIBUTE | | 106 |
| 4.1 | Introduction | 106 |
| 4.2 | Related Works | 110 |
| 4.2.1 | Access Control in EHR | 110 |
| 4.2.1.1 | RBAC in EHR Security | 110 |
| 4.2.1.2 | ABAC in EHR | 110 |
| 4.2.1.3 | CAAC in EHR | 111 |
| 4.2.2 | Ontology in EHR Security | 111 |
| 4.2.3 | Summary | 112 |
| 4.3 | Proposed Framework: GPT-Onto-CAABAC | 112 |
| 4.3.1 | High-level framework overview | 113 |
| 4.3.2 | Detailed ontology explanation | 114 |
| 4.3.3 | Detailed CAABAC explanation | 116 |
| 4.3.3.1 | Advantages of ad hoc contextual information in healthcare | 116 |
| 4.3.3.2 | Role of CAAC | 116 |
| 4.3.3.3 | Contribution of ABAC | 117 |
| 4.3.3.4 | Distinction between CAABAC and ABAC | 117 |
| 4.3.3.5 | GPT-Onto-CAABAC context capture | 117 |
| 4.3.3.6 | Example of CAABAC | 118 |
| 4.3.4 | GPT integration and conflict resolution | 118 |
| 4.3.5 | Human oversight and sign-off | 119 |
| 4.4 | Implementation of the GPT-Onto-CAABAC framework | 120 |
| 4.4.1 | Construction of policy-to-legal-ontology | 121 |
| 4.4.2 | Utilisation of datasets | 122 |
| 4.4.3 | Acquiring decisions and recommendations | 123 |
| 4.4.4 | Human evaluation and sign-off | 124 |
| 4.4.4.1 | Compliance | 124 |
| 4.4.4.2 | Adaptability | 125 |

- 4.4.4.3 Conflict Resolution Efficiency 125
- 4.4.4.4 Recommendation Quality 125
- 4.5 Evaluations 126
 - 4.5.1 Scenario Testing with Evaluation Metrics 126
 - 4.5.1.1 Scenario Testing 127
 - Contextual Comprehension: 127
 - Recommendations Effectiveness: 127
 - Overall Performance: 128
 - 4.5.1.2 Fault Injection Testing 129
 - 4.5.1.3 GPT Responses Patterns 129
 - 4.5.2 Comparative Evaluation 131
 - 4.5.3 Ethical and societal implication analysis 132
 - 4.5.4 Assessment of Transparency and Interpretability 133
- 4.6 Discussions 133
 - 4.6.1 Challenges and Overcoming Strategies 134
 - 4.6.2 Detailed Dataset Discussion 135
 - 4.6.3 Applications in Healthcare Settings 135
 - 4.6.4 Expanded use cases beyond EHR 136
 - 4.6.5 Translating concept to real-world implementation 137
 - 4.6.6 Expanded Experimental Comparison 139
 - 4.6.7 Research limitations 140
 - 4.6.8 Future research directions 140
- 4.7 Conclusions 141

PART B: ENHANCING HEALTH INFORMATION SYSTEM SECURITY: AN ONTOLOGY MODEL APPROACH 145

- 4.8 Introduction 145
 - 4.8.1 The current status of information security related to EHR 146
 - 4.8.2 The current state of research and a brief introduction of its inadequacies of security EHR 147
 - 4.8.3 The Contributions and the Outline of the Study 148
- 4.9 Research Motivations 150
 - 4.9.1 Motivating EHR Use Case Scenarios 150
 - 4.9.2 Scenario 1: Patient-Centered Care Coordination 151
 - 4.9.3 Scenario 2: Patient Data Privacy 151
 - 4.9.4 Scenario 3: Emergency Room Admission 152
 - 4.9.5 Scenario 4: Clinical Research 152
 - 4.9.6 Scenario 5: Unauthorized Access and Disclosure of Mental Health Records 153
 - 4.9.7 Scenario Analysis 153
- 4.10 Conceptual Ontology Security Model 154
 - 4.10.1 Background 154
 - 4.10.2 Ontology Conceptual Security Model 154
 - 4.10.3 Identifying Stakeholders 156
 - 4.10.4 Identifying Health Information 156
- 4.11 Security or Policy Ontology 157
 - 4.11.1 Security or Policy Ontology 157

| | | |
|----------|--|------------|
| 4.11.2 | Entities | 158 |
| 4.11.3 | Attributes | 159 |
| 4.11.4 | Operations | 160 |
| 4.11.5 | Security Conditions | 161 |
| 4.11.6 | SWRL Rules | 161 |
| 4.12 | SWRL Rule-Bases | 162 |
| 4.13 | Role-Based Access Control Approaches | 165 |
| 4.13.1 | Context-aware Role-Based Access Control Approaches | 165 |
| 4.13.2 | Other Access Control Approaches | 168 |
| 4.13.3 | Summary | 170 |
| 4.14 | Evaluation and Discussion | 172 |
| 4.14.1 | Research Limitations | 172 |
| 4.14.2 | Research Challenges | 175 |
| 4.14.3 | Summary | 178 |
| 4.15 | Related Works | 178 |
| 4.16 | Conclusion and Future Research | 182 |
| 5 | TOWARDS A UNIVERSAL PRIVACY MODEL FOR ELECTRONIC HEALTH RECORD SYSTEMS: AN ONTOLOGY AND MACHINE LEARNING APPROACH | 188 |
| 5.1 | Introduction | 188 |
| 5.2 | Related Work | 191 |
| 5.2.1 | Personally-Controlled EHR Systems | 191 |
| 5.2.2 | Ensuring Privacy through Smart Contract—Healthcare Blockchain Systems | 192 |
| 5.2.3 | Context-Sensitive Privacy Policies | 193 |
| 5.2.4 | Homomorphic Encryption in EHR Systems | 194 |
| 5.2.5 | Comparison with Our Study | 196 |
| 5.3 | Research Motivation | 197 |
| 5.3.1 | Use Case Scenarios | 199 |
| 5.3.1.1 | Scenario 1: Primary Care Physician | 200 |
| 5.3.1.2 | Scenario 2: Emergency Care | 200 |
| 5.3.1.3 | Scenario 3: Clinical Research | 200 |
| 5.3.1.4 | Scenario 4: Multidisciplinary Consultation | 201 |
| 5.3.1.5 | Scenario 5: Telehealth | 201 |
| 5.3.1.6 | Scenario 6: Data Breach | 201 |
| 5.3.2 | Research Challenges | 202 |
| 5.4 | A Privacy Model for EHR Systems | 203 |
| 5.4.1 | Leveraging Ontology and ML for Enhanced e-Healthcare Privacy | 203 |
| 5.4.1.1 | Intrusion Detection and Prevention | 203 |
| 5.4.1.2 | Confidentiality and Privacy of EHR | 203 |
| 5.4.1.3 | Improved Indexing and Retrieval Performance | 204 |
| 5.4.1.4 | Secure Data Access and Privacy Preservation | 204 |
| 5.4.1.5 | Privacy Disclosure Measurement | 204 |
| 5.4.1.6 | Efficiency and Accuracy in Clinical Information Extraction | 204 |
| 5.4.1.7 | Structured Approach to organizing Clinical Data | 204 |
| 5.4.2 | Conceptual Privacy Models | 205 |

| | | |
|----------|---|------------|
| 5.4.3 | Identifying Stakeholders | 205 |
| 5.4.4 | Redefining Health Information and Privacy Rules | 206 |
| 5.4.5 | Role Ontology | 210 |
| 5.4.6 | Health Information Ontology | 211 |
| 5.4.7 | Privacy Policy Ontology | 211 |
| 5.4.8 | Disclosing Emergency Health Information for Patients in Car Accident Case Study | 211 |
| 5.5 | Evaluation of Privacy Ontology and Experiments | 213 |
| 5.5.1 | Experiment Setup and Dataset Preparation | 214 |
| 5.5.2 | Clarifying Privacy Policy Classification | 215 |
| 5.5.3 | Development of ML-Based Privacy Model | 216 |
| 5.5.4 | Application of Privacy Ontology in the Machine Learning Experiment | 218 |
| 5.5.5 | Evaluation Results | 219 |
| 5.5.5.1 | Technical Details | 219 |
| 5.5.5.2 | Dataset and Results | 219 |
| 5.5.6 | Summary of the Findings | 220 |
| 5.5.7 | Refined Technical Approach and Dataset Overview | 221 |
| 5.5.7.1 | In-depth Technical Methodology | 221 |
| 5.5.7.2 | Comprehensive Dataset Description | 222 |
| 5.5.7.3 | Ensuring Reproducibility | 222 |
| 5.6 | Discussion | 222 |
| 5.6.1 | The Relationship between the Ontology and the ML Model | 223 |
| 5.6.2 | The Complementarity of Ontology and ML Model | 223 |
| 5.6.3 | Adoption of Privacy-Preserving Technologies for Health Information Security | 224 |
| 5.6.4 | Section Summary | 224 |
| 5.7 | Conclusions and Future Research | 224 |
| 5.7.1 | Limitation | 225 |
| 5.7.2 | Future Research Directions | 225 |
| 6 | CONCLUSION | 227 |
| 6.1 | Introduction | 227 |
| 6.2 | Thesis Contributions | 228 |
| 6.2.1 | Mapping Thesis Contributions to Research Questions | 230 |
| 6.2.2 | Major Contributions to the Field | 232 |
| 6.2.3 | Implications for Practice and Policy | 232 |
| 6.2.4 | Study Limitations | 233 |
| 6.3 | Future Research Directions | 234 |
| 6.3.1 | Advancing CEMPS with Emerging Technologies | 235 |
| 6.3.2 | Ontology and Machine Learning Synergies | 235 |
| 6.3.3 | Blockchain for Decentralized Trust | 235 |
| 6.3.4 | Ethical, Legal, and Social Implications (ELSI) | 235 |
| 6.3.5 | Interoperability and Standardization | 236 |
| 6.3.6 | User-Centered Design and Usability | 236 |
| 6.3.7 | Comprehensive Evaluation and Real-World Implementation | 236 |

Bibliography

237

List of Tables

| | | |
|-----|--|-----|
| 1.1 | Advanced Analysis of EHR Complications and Strategic Solutions | 8 |
| 1.2 | Core Principles of CIA in EHR | 11 |
| 1.3 | Comparison of Centralized vs. Decentralized EHR Systems | 14 |
| 2.1 | PCS: Privacy, Confidentiality, and Security. | 29 |
| 2.2 | List of Target Journals and Conferences | 37 |
| 2.3 | EHR Complications and Potential Solutions | 49 |
| 2.4 | Clarifying the Components of Information Security | 50 |
| 2.5 | Categorization of Technologies, ABC (Access Control, Blockchain, Cloud-based, Cryptography) | 53 |
| 2.6 | Comparative Analysis of Blockchain-based Solutions for EHR | 60 |
| 2.7 | Difference Between Privacy, Confidentiality and Security | 64 |
| 2.8 | Summary Of Reviewed Technologies And The Aspects They Covered | 68 |
| 3.1 | Comparative Analysis of Centralized and Decentralized EHR Models | 97 |
| 3.2 | Comparative Analysis of Privacy Preservation Techniques in Healthcare | 100 |
| 4.1 | Comparison of different access control models in addressing extrinsic and intrinsic factors (✓: capable; △: partially capable; ×: incapable) | 113 |
| 4.2 | List of legislations governing EHR access | 121 |
| 4.3 | Marking rubric for evaluating GPT responses | 126 |
| 4.4 | MEDICATION ADMINISTRATION PATTERN TRIPLE-BASED REPRESENTATION | 149 |
| 4.5 | Data Quality Ontology – Key Concepts in HER Applications | 155 |
| 4.6 | Comparison of Access Control Models in Healthcare Settings | 171 |
| 5.1 | List of role-based stakeholders and privacy rules. | 199 |
| 5.2 | The modeling criteria of the ontology. | 208 |
| 5.3 | A case study of a health emergency. | 213 |
| 5.4 | An example privacy policy. | 217 |
| 5.5 | Experiment results using NLP-based BERT techniques. | 220 |

List of Figures

| | | |
|------|--|-----|
| 1.1 | Interconnection between Problem Statement, Research Questions (RQ1-RQ5), and Thesis Contributions | 18 |
| 2.1 | The Cyber Kill Chain illustrating the points of compromise for Security, Privacy, and Confidentiality in EHR systems | 33 |
| 2.2 | Numbers and Percentages of Publication Types | 38 |
| 2.3 | Distribution of Research Articles from 2013 to 2022 | 53 |
| 2.4 | Number of (Reviewed) Technologies Covering the Aspect(s) i.e., Privacy, Confidentiality, Security | 65 |
| 3.1 | EHR Generic Architecture | 77 |
| 3.2 | Enhanced EHR Integration Program | 80 |
| 3.3 | Architectural Framework of CEMPS | 81 |
| 3.4 | Simulating Local Data for Each Node | 83 |
| 3.5 | Defining FL data storage model for CEMPS | 84 |
| 3.6 | Federated Averaging and Model Training | 84 |
| 3.7 | Data Encryption - Cryptography Libraries | 85 |
| 3.8 | Integrate Encryption | 85 |
| 3.9 | Data Sharing and Access Layer with DP | 88 |
| 3.10 | Various Stages of defining the Privacy Model | 92 |
| 3.11 | Architectural Framework of CEMPS | 94 |
| 4.1 | Number of Larger Data Breaches (≥ 500 Records Per Breach) of EHR from 2009 to 2022 in USA | 108 |
| 4.2 | GPT-Onto-CAABAC | 113 |
| 4.3 | Evaluation of GPT Answers Per Category (higher is better) | 128 |
| 4.4 | Variation of Evaluation Scores of GPT Responses By Category | 130 |
| 4.5 | Comparison of our GPT-4-based prototype (left) and a practical domain knowledge LLM implementation (right) | 137 |
| 4.6 | GPT-Onto-CAABAC Implementation Roadmap for 2024 | 139 |
| 4.7 | Knowledge structure of EHR triplestore. | 148 |
| 4.8 | Creating an SWRL rule-base | 163 |
| 4.9 | System Interaction | 175 |
| 4.10 | Semantic Middleware Architecture. | 179 |
| 5.1 | The relevant concepts to build the privacy model. | 207 |
| 5.2 | The core concepts of privacy ontology. | 208 |
| 5.3 | Role ontology. | 209 |
| 5.4 | Health information ontology. | 209 |

5.5 Privacy policy ontology. 209

List of Abbreviations

| | |
|---------------|--|
| <i>ABAC</i> | Attribute-based Access Control |
| <i>AIS</i> | Artificially Intelligent Systems |
| <i>AI</i> | Artificial Intelligence |
| <i>APP</i> | Australian Privacy Principles |
| <i>CAABAC</i> | Context-Aware Attribute-Based Access Control |
| <i>CAAC</i> | Context-Aware Access Control |
| <i>COPD</i> | Chronic Obstructive Pulmonary Disease |
| <i>CPPM</i> | Centralized Privacy-Preserving Model |
| <i>DL</i> | Description Logic |
| <i>EDI</i> | Electronic Data Interchange |
| <i>EDs</i> | Emergency Departments |
| <i>EHR</i> | Electronic Health Record |
| <i>EMR</i> | Electronic Medical Record |
| <i>FHE</i> | Fully Homomorphic Encryption |
| <i>GDPR</i> | General Data Protection Regulation |
| <i>GPT</i> | Generative Pre-trained Transformers |
| <i>HDB</i> | Hippocratic Databases |
| <i>HIPAA</i> | Health Insurance Portability and Accountability Act |
| <i>HITECH</i> | Health Information Technology for Economic and Clinical Health |
| <i>HIS</i> | Health Information Systems |
| <i>MAC</i> | Mandatory Access Control |
| <i>MHR</i> | My Health Record |
| <i>NFR</i> | Not For Resuscitation |
| <i>NLP</i> | Natural Language Processing |
| <i>PHR</i> | Personal Health Record |
| <i>PII</i> | Personally Identifiable Information |
| <i>PKI</i> | Public Key Infrastructure |
| <i>RBAC</i> | Role-Based Access Control |
| <i>UKNHS</i> | United Kingdom National Health System |

Chapter 1

INTRODUCTION

This chapter provides an introduction to the thesis topic and context of the research presented in this thesis, namely, A Security and Privacy Compliant Data Sharing Solution For Healthcare Data Ecosystem.

The growing use of electronic health-related data has ignited a research interest that spans diverse domains, industries, and stakeholders. This transformative evolution of data has enabled the healthcare industry to convert health data into electronic health records (EHR) or electronic health records. EHR encompasses electronic patient records that contain demographic information, medical histories, medication records, allergies, immunization status, laboratory results, radiology images, billing details, etc. The adoption of **EHRs** offers numerous advantages, including rapid access to clinical data, streamlined clinical workflows, error reduction, improved patient safety, cost savings, and improved support for clinical decision making. EHRs facilitate the creation, storage, management, and access on demand of health information for both healthcare providers and patients. In this context, various services such as cloud services and blockchain have emerged as a robust infrastructure, reducing the cost of data storage processing and maintenance while enhancing efficiency and data quality. Regarding data centralization, the vast network of remote servers, although accessible from multiple locations, introduces security and privacy challenges, particularly given the sensitive and confidential nature of medical data. Security and privacy concerns within EHRs are multifaceted. They encompass the potential inference of private information when combined with external datasets, the exploitation of user data to benefit organizations, social stratification based on data literacy, and the risk of adverse consequences without awareness or defense mechanisms. The digital divide prevalent in some regions intensifies these challenges. The significance of EHR security cannot be overstated, particularly given the widespread adoption of EHR systems, the increasing cybersecurity threats, the critical role that EHRs play in patient care, and the ethical and legal obligations to protect patient data. The quest for

a robust privacy-preserving health data sharing framework within the healthcare sector has been propelled by the need to preserve the privacy of its users. The Australian healthcare authorities have expressed concerns about the existing My Health Record (MHR) system, calling for improvements due to data breach incidents. In particular, the healthcare sector has become a significant focus for identity fraud because health records also include private information such as patient identities, credit card numbers, and addresses. Additionally, developments in information and communication technology have resulted in a situation in which patient health records pose new protection and privacy risks. Hackers, viruses, and worms can seriously endanger the protection and privacy of EHRs. Concerns about data protection and privacy have arisen, according to research conducted in many countries. According to new surveys [1–5], it is estimated that there are at least 25 million statutory authorizations executed each year in the United States for the release of health information. The participants mentioned [6] their concerns about the protection of the EHR data in all surveys conducted in Denmark, Germany, and New Zealand. As stated above, the main and most serious hurdle to the implementation of EHRs are privacy and security issues. Although there are various compliance techniques that can be introduced to deter unwanted entry into electronic health information, depending on the scale and nature of a healthcare institution, it is impossible to determine with certainty what procedures can and should not be implemented. The *Health Insurance Portability and Accountability Act* (HIPAA) and the *Health Information Technology for Economic and Clinical Health* (HITECH) Act have already implemented many safety protections in the EHR [7]. However, in existing security procedures, HIPAA remains behind from a security perspective, and data encryption is considered an addressable requirement. It does not describe the ways in which the frameworks now being used will be developed and enforced. This contributes to many variations in fragmented processes today and has discouraged interoperability between medical institutions [8]. Previously, the use of blockchain technologies for the provision of safe and stable health records, the exchange of biomedical and electronic health data, brain modeling and reasoning has been of great concern [9]. Blockchain technology generally has the key characteristics of decentralization, persistence, anonymity, and auditability. With these characteristics, the blockchain can save a lot of money and improve efficiency. Therefore, in the era of virtualization, data privacy is being protected using various traditional techniques, such as encryption and blockchain. In the recent literature, researchers have highlighted newly evolved techniques and tools to achieve privacy, security, and authentication of electronically stored health data [10, 11].

1.1 Comparative Advantages of EHRs

EHRs comprise patient data in digital form that are stored and exchanged securely and are accessible by multiple authorized users to support the continuous and efficient management of integrated healthcare [12–14]. These electronic data are processed quickly and can be sent to digital devices. Technically, EHRs are designed with the purpose of constantly providing and delivering reliable data to health organizations. It comprises details on the patient’s medical history, including diagnosis, laboratory findings, information on hospital admissions, treatments, surgical procedures, and medications with the least mistakes, more effectiveness, and better care. They even describe the patient’s condition, allowing for a more detailed diagnosis and treatment of the patient [15]. EHRs can be shared with other healthcare providers when necessary. However, EHRs are prone to various types of security and privacy attacks during transmission [15]. In distributed medical research and healthcare systems, the assurance of data privacy is based on compliance with laws and jurisdictions, as highlighted by various studies [16–18]. Although there is a compelling need to enforce privacy policies at the program level because existing measures often do not provide sufficient guarantees for effective privacy protection [19]. To address this gap, it becomes crucial to critically examine privacy agreements, which contribute to the overall improvement of social acceptance in healthcare systems. Incorporating health mechanisms is a viable approach, particularly in Australia, offering enhanced privacy features. This, in turn, empowers people throughout the country to access superior health and medical services with less concern about the confidentiality of their data [20]. Therefore, there is a pressing need for the implementation of robust mechanisms that can support the efficacy of privacy protocols and processes within patient data management systems. This requires a proactive approach that includes auditing and monitoring of past data sharing practices in the healthcare system. In light of its widespread use, the development of a safe environment for the sharing of EHRs has gained a great deal of interest in the healthcare sector. The most recent literature [21–24] indicates that there are many benefits to using EHR software, including cost savings, increased quality of healthcare, advancement of evidence-based medicine, more comprehensive data collection, and flexibility. Consequently, the term EHR in this thesis refers not only to an electronic database to store and retrieve health information, but also to a system that can be used to enforce and maintain completeness of data, resilience to failure, high availability, and consistency of security policies. Finally, this work acknowledges the different nature of health information such as *Personal Health Records* (PHR) and *Electronic Medical Records* (EMR) and their privacy, but this thesis focuses specifically on EHR.

1.1.1 Electronic Health Records Benefits

One of the key advantages of an EHR is that it allows health data to be structured and stored in a digital design that supports exchange with multiple external organizations that provide support to the primary health institution by certified providers. In 2018, a study by IBM and the Ponemon Institute revealed that health care violations are the most expensive among industries such as trade, finance, government, etc., based on an average cost of 380 dollars per record. Many current EHR programs are quite often unreliable, resulting in poor patient care and poor interoperability with other systems. Compared to the speed and complexity of the records produced, the development of these applications remains low [8]. The development of a safe environment for sharing EHR has strained a lot of interest in the healthcare sector with the widespread use of EHR [25]. The terms EMR and EHR sometimes cause confusion. *Electronic Medical Records* (EMR) contain medical and clinical data collected from the provider's office, while an EHR includes more comprehensive patient information [26]. An EHR is a comprehensive record of clinical and administrative data on all individuals cared for within healthcare systems. As such, it offers a repository of sufficient size and scope to support detailed clinical care analysis and evaluation of important subgroups of patients, such as those with serious disease, including "high-cost, high-need patients" who will particularly benefit from palliative care [27–31]. Additionally, EHR data is recorded directly from documentation of healthcare care delivery and can be used without requiring additional data entry beyond accompanying routine care. EHRs also capture social determinants of health, which are important for the provision of responsible high-quality care to seriously ill [27, 32]. There are many benefits to EHRs, for example, higher care levels, improved patient safety, simplified chaotic processes, and reduced costs. The portability and advantages of having digitized records that can be accessed and used anywhere during the clinical decision-making process promised to herald a new era in patient care [32]. With the ability to address public health and population information needs, it can contribute to the creation of health policies, decision making, and the promotion of healthier lifestyles [33]. An EHR can be used continuously to improve communication, improve quality of care, reduce medical errors, and reduce waste [34]. It also has the potential to transform the healthcare system from a mostly paper-based industry to one that uses multiple sets of information to help providers provide higher quality of care to their patients [35]. EHRs accelerate information access and have the potential to improve clinical workflow; they also have the capacity to support other associated activities using various tools such as *Decision Support System* (DSS) and *Intelligent Systems* (IS) [36]. The *Health Information Technology for Economic and Clinical Health Act* (HITECH Act) was introduced in the United States in 2009 to facilitate and accelerate the adoption of EHR and supporting technology [38]. The government invested more than 30 billion dollars and 95%

hospitals implemented the EHR, which was a significant increase from only 9% in 2008 [37]. The HITECH Act rewards hospitals and medical professionals for adopting EHRs that aim to improve the quality of care in the NHS. The key benefits identified were: 1. Greater convenience through structured information access, 2. Better communication to improve patient-related outcomes, 3. Faster evidence-based decision making during care delivery and 4. Optimization of resources through effective use and productivity at all levels [32, 37, 38]. EHRs have provided many benefits to all stakeholders at various levels. For example, it allows for early diagnosis of diseases, has reduced medication errors (drug overdoses, adverse effects, drug interactions, allergic reactions *etc.*) by 95%, has improved lifesaving measures by ensuring compliance with care adherence, and has decreased the number of duplicate diagnostic tests and lowered costs by 7-11% [39, 40]. There are other specific advantages: management of epidemics in developing countries, better informed decision-making, care coordination, patient satisfaction and better outcomes, and improvement of care based on real-world evidence [32]. Therefore, the EHR offers important opportunities to identify seriously ill patients and assess palliative care quality metrics in a large number of patients with serious disease. In an EHR SWOT analysis, the highest priority in the strength analysis was timely and quick access to information and the ability to store information, as confirmed in [41].

1.2 Synergistic Integration of Technologies for EHR Enhancement

The advent of Electronic Health Records (EHRs) has ushered in a new era of healthcare, where information is digitized, leading to improved healthcare delivery and patient outcomes. However, the digitization of health information also introduces significant challenges, particularly in the realms of security and privacy. Addressing these challenges necessitates a multifaceted approach, integrating advanced technologies such as Access Control, Blockchain, Cloud Computing, and Cryptography. This section elucidates the synergies among these technologies and their collective contribution to fortifying EHRs against security and privacy breaches.

Access Control mechanisms serve as the first line of defense, ensuring that only authorized users gain access to sensitive health information [6, 7, 42, 43]. By delineating clear access rights based on roles or attributes, Access Control systems prevent unauthorized access, a fundamental aspect of preserving privacy and integrity within EHRs.

Blockchain technology, with its inherent characteristics of decentralization, immutability, and transparency, offers a robust framework for the secure sharing and storage of

EHRs [9]. By creating a tamper-proof ledger of health records, Blockchain ensures the integrity and non-repudiation of health data, facilitating a trustless environment where stakeholders can share information securely.

Cloud Computing provides a scalable and efficient infrastructure for storing and processing the vast amounts of data generated by EHR systems [10]. The elasticity of cloud services allows healthcare providers to manage resource demands dynamically, ensuring availability and accessibility of health information. Furthermore, cloud platforms can leverage advanced security and privacy features, offering a secure environment for hosting EHRs.

Cryptography plays a pivotal role in securing data at rest and in transit, protecting against eavesdropping and unauthorized access [8]. Through encryption, Cryptography ensures that health data remains confidential, safeguarding patient privacy. Additionally, cryptographic techniques enable secure authentication and verification processes, further enhancing the security posture of EHR systems.

Integrating these technologies provides a comprehensive security and privacy-preserving solution for EHRs. Access Control restricts access, Blockchain ensures data integrity and trust, Cloud Computing offers scalability and resilience, and Cryptography secures data confidentiality. Together, they form a synergistic framework that addresses the multifaceted challenges faced by EHRs, paving the way for a secure, efficient, and privacy-compliant healthcare data ecosystem.

1.2.1 EHR Drawbacks

Despite huge expectations and investment, EHRs have not been entirely successful in addressing what they were established to rectify. Fundamental problems started at the initial execution stage. There were issues in feeding and safeguarding patient clinical information, hospital inventories, staffing, and resources when computing devices started to be used extensively, there were cost issues and a lower than expected return on investment, technical failures, privacy and confidentiality concerns, and a lack of resources *that is*, hardware and infrastructure [44]. The key identified weaknesses are: lack of harmony, problems in patient matching, data security and privacy concerns, and algorithm manipulation in decision support models that contribute to clinician burden, redundant credentials, EHR workflows with many phases and complexity which require automation using modern technologies, and insecure data storage that enables inappropriate use by end users [32]. In addition to these flaws, EHRs also have a few drawbacks (**Table 1.1**). For example, data acquired from the EHRs represent treatment that was documented rather than actual care that was provided or how patients and their families

experienced it. A good illustration of this possible discrepancy is the goal-of-care talk, which physicians may claim to have had with patients, but neglect to record it or they record it in a nonretrievable manner. The validity of EHR data for use to ensure quality and liability can be compromised by the wide range of objectives for which they are collected. Another drawback of EHRs is the lack of critical outcomes necessary to provide high-quality disease treatment. Therefore, EHRs manage and store patient medical records in various medical facilities in a centralized computer database. Assessment of patient risk is facilitated by EHR through the analysis of medical reports. Furthermore, these systems facilitate the exchange of patient data and information over the Internet, allowing physicians to diagnose and treat patients seamlessly between different medical institutions [45]. Despite the advantages, EHRs exhibit security vulnerabilities, particularly during data transmission over the Internet or data retrieval from a server database, and lack systematic and organized evaluation of results in general [27]. Consequently, the integrity of EHR systems is highly dependent on robust security features.

1.3 Confidentiality, Integrity, and Availability (CIA) in EHR

Confidentiality, Integrity, and Availability (CIA) constitute the fundamental principles of information security within EHR systems. These principles collectively safeguard patient information, ensuring its protection, accuracy, and accessibility. In the context of EHRs, each facet of the CIA triad plays a critical role in maintaining the security and reliability of healthcare data. The principles of Confidentiality, Integrity, and Availability, collectively known as the CIA triad, are foundational to securing information systems. In the context of Electronic Health Records (EHRs), these principles take on heightened significance due to the sensitivity of health information and the potential impacts on patient care and privacy.

Confidentiality in EHR systems is paramount to maintaining patient trust and ensuring compliance with stringent regulations such as the Health Insurance Portability and Accountability Act (HIPAA) [46, 47]. Confidentiality ensures that patient data, from diagnoses to treatment plans, is accessible only to authorized individuals. Recent advances in encryption technologies have strengthened confidentiality in EHRs. For example, end-to-end encryption (E2EE) ensures that patient data transmitted over networks cannot be intercepted and read by unauthorized parties [8]. Despite these advances, the proliferation of ransomware attacks, as discussed in recent statistics on cyber threats to healthcare systems, presents ongoing challenges to maintaining confidentiality [48–50].

TABLE 1.1: Advanced Analysis of EHR Complications and Strategic Solutions

| EHR Complication | Strategic Solution | Technological/ Methodological Implementation |
|----------------------------------|---|--|
| Redundant Documentation | Streamline Clinical Documentation | Utilize AI-driven data capture and analysis to identify and include only essential clinical information. |
| Complex Workflow | Simplify and Optimize Workflow | Implement process mining techniques to identify inefficiencies and redesign the workflow for optimal data entry and retrieval. |
| Need for Advanced Automation | Enhanced Automation Integration | Leverage advanced voice recognition and natural language processing (NLP) to facilitate hands-free data entry and retrieval. |
| Proprietary Software Limitations | Open Platform Development | Foster an ecosystem of open source EHR platforms, encouraging innovation and customization. |
| Data Silos | Seamless Data Exchange and Interoperability | Adopt FHIR (Fast Healthcare Interoperability Resources) standards and smart APIs to enhance data sharing and interoperability. |
| Suboptimal User Experience | User-Centric Design | Apply UX/UI design principles, including mobile-first strategies, to improve accessibility and ease of use. |
| Security Vulnerabilities | Robust Security Measures | Integrate advanced cybersecurity protocols, including blockchain for data integrity and AI-based anomaly detection systems. |
| Regulatory Compliance | Compliance Automation | Implement regulatory compliance management software with real-time updates on policy changes and compliance tracking. |
| Data Analytics | Enhanced Analytic Capabilities | Use big data analytics and machine learning algorithms for predictive analytics and decision support. |
| Patient Engagement | Interactive Patient Portals | Develop User-Friendly Patient Portals with personalized health tracking, telehealth, and secure messaging features. |

Integrity involves ensuring the accuracy and completeness of patient data within EHRs. It is crucial to provide high-quality healthcare, as even minor errors or alterations in health records can lead to misdiagnosis or inappropriate treatment. Blockchain technology has emerged as a promising tool for improving data integrity in EHR systems. By providing an immutable ledger of health records, blockchain technology ensures that once patient data is recorded, it cannot be altered or deleted without detection [9]. However, challenges remain in integrating blockchain with existing healthcare IT systems and ensuring scalability.

Availability is critical to ensure that healthcare providers have timely access to EHRs for effective patient care, especially in emergency situations. Cloud computing offers a solution to improve the availability of EHRs using distributed resources to ensure redundancy and resilience against system failures [10]. However, the reliance on cloud platforms introduces challenges such as potential downtimes and the risk of Distributed Denial of Service (DDoS) attacks, which can temporarily render EHR systems inaccessible.

In conclusion, while significant advances in technologies such as encryption, blockchain, and cloud computing have bolstered the CIA triad in EHR systems, these advancements are not without new challenges. The continuous evolution of cyber threats requires ongoing vigilance and innovation in applying the principles of the CIA triad to safeguard EHRs. The statistics on cyber attacks discussed earlier underscore the importance of robust security measures in protecting healthcare data against evolving threats.

1.3.1 Prevalence and Impact of Cyber Attacks on Healthcare Systems

The escalation of cyber attacks on healthcare systems has become a critical concern, with Electronic Health Records (EHRs) emerging as a primary target due to the sensitive nature of the data they contain. Recent statistics illuminate the severity and frequency of these breaches. For example, a report by the Department of Health and Human Services (HHS) indicates that healthcare care breaches have affected more than 26 million individuals in the United States alone, as of the last reporting year [51–53]. Furthermore, the cybersecurity firm Cybersecurity Ventures predicts that the costs associated with ransomware damage will exceed \$20 billion worldwide by 2021, with healthcare systems being the most frequently targeted entities [54–56].

EHR systems, in particular, face unique vulnerabilities as they store comprehensive patient information, making them attractive targets for cybercriminals. The Cyber Kill Chain, investigated in Chapter 2, delineates the various stages of a cyber attack, providing information on potential security breaches within EHR systems. It emphasizes the critical points of compromise that could affect Security, Privacy, and Confidentiality, underscoring the importance of implementing robust security measures to mitigate these risks [57–60].

These statistics and examples underscore the urgent need for enhanced security protocols and innovative solutions within healthcare IT infrastructures. The vulnerabilities exposed by these attacks not only jeopardize patient privacy and data integrity, but also highlight the potential operational and financial consequences for healthcare providers.

Consequently, the research and methodologies proposed in this thesis aim to address these significant challenges by advancing the security and privacy capabilities of EHR systems, thus mitigating the risks associated with cyber threats in the healthcare sector[61–70].

1.3.2 Comparative Analysis of Global Healthcare Data Policies

The landscape of global healthcare data policies is marked by significant diversity, with frameworks such as the Health Insurance Portability and Accountability Act (HIPAA) in the United States and the General Data Protection Regulation (GDPR) in Europe serving as key benchmarks. These policies aim to safeguard patient data, although through different approaches and emphases, which directly influence the management and security of Electronic Health Records (EHRs).

HIPAA provides a solid foundation for privacy and security in the US healthcare system, focusing on the protection of personal health information. The HIPAA Security Rule mandates specific physical, administrative and technical safeguards for EHRs, including access controls and audit trails [47]. However, the applicability of HIPAA is primarily within the US, posing challenges for international data exchanges and collaborations in healthcare.

GDPR, on the other hand, offers a broader protection scope, applicable to any entity processing personal data of EU residents. GDPR emphasizes the rights of the data subject, including the right to access, rectify, and erase personal data, which extends to health records [71–73]. Although GDPR provides a stringent framework for data protection, its rigorous consent requirements pose challenges for seamless integration and use of EHRs in different healthcare services.

Comparative Challenges include the interoperability between systems governed by different regulatory frameworks. For example, the transfer of patient data between entities covered by HIPAA and those under GDPR requires careful navigation of both sets of regulations, often necessitating additional measures to ensure compliance on both fronts. Furthermore, both frameworks face challenges in keeping pace with the rapid advances in digital health technologies, such as telemedicine and mobile health apps, which may fall outside the traditional bounds of EHR systems, yet have significant implications for patient privacy and data security.

In conclusion, while HIPAA and GDPR represent critical efforts to protect healthcare data, the disparities and gaps between these and other global policies highlight the complexities of managing EHRs in a globalized world. The ongoing evolution of healthcare technologies and the international flow of health data call for dynamic, interoperable policy frameworks that can adapt to new challenges while ensuring the confidentiality, integrity, and availability of EHRs.

1.3.3 Academic Perspectives in EHR

Academic research in the field of EHR security underscores the critical importance of the CIA triad. Studies, such as Magnus et al. (2020) [74], emphasize the importance of confidentiality, integrity, and availability in protecting patient information. These studies explore the implementation of technical safeguards, such as encryption and access controls, as effective measures to preserve the CIA triad.

Frank et al. (2018) [21] investigate the role of administrative procedures, including policies and training programs, in maintaining the CIA triad. They highlight the need for a comprehensive approach that combines both technical and administrative safeguards to ensure the security of EHR systems.

In conclusion, the CIA triad serves as the basis for information security in EHRs, protecting patient data, accuracy, and accessibility. Academic research in this domain underscores the critical role of confidentiality, integrity, and availability and highlights the multifaceted approach required to uphold these principles effectively. Protecting the CIA triad is not only technologically important, but is also essential in building patient trust and ensuring the delivery of high-quality healthcare.

In this section, we present a table (Table 1.2) that summarizes and explains the core principles of Confidentiality, Integrity, and Availability (CIA) in EHR systems. The CIA triad is fundamental to ensuring the security and reliability of patient information within EHRs. Each aspect of the triad plays a crucial role in safeguarding healthcare data.

TABLE 1.2: Core Principles of CIA in EHR

| Principle | Description | Implementation Measures |
|------------------------|--|--|
| Confidentiality | Protection of sensitive patient information from unauthorized access, use or disclosure. | <ul style="list-style-type: none"> - Access controls - Data encryption - Audit logs |
| Integrity | Assurance that patient information is accurate, complete, and unaltered. | <ul style="list-style-type: none"> - Data validation rules - Digital signatures - Data backups |
| Availability | Ensuring patient information is accessible to authorized personnel when needed. | <ul style="list-style-type: none"> - Redundancy and fault tolerance - High availability configurations - Regular system maintenance |

1.4 Privacy Preservation of EHR

EHRs contain sensitive and personal health information; therefore, protecting patient privacy is crucial. Privacy in the EHR involves the implementation of principles, policies, and measures that protect the confidentiality of patient information and prevent unauthorized access, use, or disclosure. Technical safeguards such as encryption, authentication, access controls, and audit logs are essential to protect the privacy of the EHR [52]. Administrative procedures, such as policies and practices that govern access, use, and disclosure of patient information, are also critical components of EHR privacy protection [75]. Patient rights play an essential role in EHR privacy and include the right to control your information, access to your medical records, and the right to file complaints if your privacy rights have been violated [76]. Therefore, protecting privacy in the EHR is vital to maintaining confidentiality, integrity, and availability of sensitive patient health information. Technical safeguards, administrative procedures, and recognition of patient rights are all necessary components to achieve this goal. Healthcare providers must ensure that all necessary measures are in place to protect patient privacy and comply with regulatory requirements. According to [77], patient personal information is a valuable asset and is considered highly sensitive and a safety net to prevent exposure to patient health. [78] indicated that data breaches that have occurred in the past have led to a variety of issues, including the disclosure of patient data, loss of credentials, malware infections, among others. Although EHRs face several access issues in the healthcare industry, data privacy and security remain the most significant issues. [79] proclaimed that newer technology models are based on a new robust and secure technology infrastructure and therefore have been more crucial for healthcare providers and organizations in maintaining patient records confidentiality. Furthermore, the process of documenting, sharing, and storing healthcare data should be carried out with the assurance of authentication and taking into account all ethical aspects around it. Although the use of digital platforms and advanced technologies for the exchange of patient information amplifies the integration of advanced technology, the challenges related to privacy, confidentiality, security, and integrity should be addressed before adopting any technology. Medical records containing personal and sensitive data are very likely to be targeted for cybercrime attacks. Their storage on centralized servers owned by other parties inadvertently raises security and privacy flaws, opening the door to a number of attacks such as ransomware [80] and DDoS [81] attacks that have more serious consequences than only financial or privacy breaches. Hackers in the United States hacked [82] into the Community Health Systems (CHS) database of a well-known hospital group, gaining access to a vast amount of private health data, including the social security numbers of more than

a million patients. Similarly to this, a DDoS attack on the websites of many hospitals, severely impairing medical services [83].

Data integrity services should ensure that patient privacy is protected not only from external hackers, but also internally within the ecosystem from rogue employees or through cloud providers [84]. Fabian, Ermakova and Junghanns [85] in the healthcare industry state that the use of technology, for example Blockchain, has raised obvious problems of privacy, safety and patient consent and brings with them the highest risks to privacy and security. The Russian Ministry of Health in partnership with *Vnesheconombank* (VEB). On 6 September 2017, as part of a health session at the *Eastern Economic Forum* (WEF), the head of the Ministry of Health said that blockchain technology could be used to store *electronic medical records* (EMR) of patients. The Ministry further emphasized that the medical record storage system will be depersonalized to the maximum extent, allowing the preservation of medical privacy. This will allow patients to determine for themselves what part of the information in their medical records they want to disclose. Gordon and Catalini [86] conducted research stating that a trend of patient-based interoperability has emerged through the conventional *electronic data exchange* (EDI) mechanism used in healthcare organizations such as hospitals and laboratories. Although patient interoperability was convenient, it resulted in privacy and security concerns due to additional access to patient data. Therefore, with respect to the security of patient data and the implementation of policies related to patient privacy, it is crucial to understand exactly what the difference is between these terms and then to ensure effective sharing of patient health data among healthcare professionals and to focus on the implementation of the explicit and controlled framework.

1.5 Centralised vs Decentralized EHR System

Despite the potential benefits of decentralized EHR systems, the literature has highlighted several critical drawbacks. For example, a study notes the challenges in ensuring consistent data quality and integrity between decentralized nodes [87]. Another significant issue is the difficulty in managing and standardizing privacy policies and access controls in a distributed environment, as explored by Smith et al. [88]. This can lead to inconsistencies in protecting patient data and potential vulnerabilities. Furthermore, Jones et al. have identified scalability and performance problems in decentralized EHR systems, especially when dealing with large-scale data and numerous stakeholders [89]. In a decentralized setup, every node in the network needs to be updated simultaneously, which can be resource intensive and may lead to inefficiencies. Furthermore, the risk of fragmentation and the lack of interoperability is a concern highlighted in the research by

TABLE 1.3: Comparison of Centralized vs. Decentralized EHR Systems

| | Characteristics | Advantages | Challenges |
|---------------------------------|---|---|---|
| Centralised EHR System | <ul style="list-style-type: none"> - Single data storage - Centralised database - Managed by a single entity - Standardised protocols and formats - Streamlined data uniformity | <ul style="list-style-type: none"> - Enhanced data uniformity, reducing data discrepancies - Simplified central management and maintenance - Efficient resource utilisation, reducing infrastructure duplication - Seamless integration with standardized formats | <ul style="list-style-type: none"> - Vulnerable to single point of failure; system downtime can affect multiple users - Attractive target for large-scale cyber attacks, necessitating robust security measures - Limited flexibility in adapting to specific requirements, potentially hindering innovation |
| Decentralised EHR System | <ul style="list-style-type: none"> - Distributed data across multiple entities - Managed by various entities - Each entity may use different protocols and formats - Diverse data sources | <ul style="list-style-type: none"> - Improved privacy and security through data distribution - Local control and customisation based on entity-specific needs - Reduced vulnerability to single points of failure, ensuring data availability | <ul style="list-style-type: none"> - Data interoperability and standardisation challenges can hinder data sharing - Complex management and higher maintenance costs, as each entity maintains its infrastructure - Ensuring consistent data quality and compliance can be challenging, requiring coordination among entities |

Doe [90]. In decentralized systems, there is a risk of creating data silos where information is not seamlessly shared across different platforms or regions, impeding the holistic view of a patient’s medical history. Furthermore, the work by Lee et al. [91] emphasizes the security risks inherent in decentralized EHR systems. The distributed nature of these systems can make them more susceptible to cyber-attacks, as multiple access points need to be secured. In Table 1.3, we provide an overview of the characteristics, advantages, and challenges associated with centralized and decentralized EHR systems. This comparison aims to highlight the key considerations when choosing between these two EHR system architectures.

This research aims to dissect and clarify the relative merits and drawbacks of centralized EHR systems in terms of security, privacy, and operational effectiveness (Table 1.3). The challenges and uncertainties associated with the deployment and management of such systems are not merely theoretical concerns, but have real-world implications for the quality of healthcare services, patient trust, and the efficiency of healthcare providers. Addressing these issues is crucial to advance the healthcare sector toward a future where secure, private, and efficient management of patient information is not an aspiration, but a reality. The resolution of these problems will not only contribute to academic knowledge, but will also have a profound impact on shaping the future of healthcare, influencing the health outcomes of individuals and communities in general.

1.5.1 Rationale for Selecting a Centralized EHR System

A centralized EHR system architecture has its challenges, especially in terms of potential vulnerability to widespread system failures or attacks, but its benefits in uniformity, efficient resource use, and compliance make it a strong candidate for healthcare data systems aiming for streamlined operations and consistent patient care.

-
- *Streamlined Data Management:* Centralised EHR systems enable more efficient management of health records, with standardised procedures for data entry, storage and retrieval.
 - *Better Resource Allocation:* Centralization can lead to cost savings due to economies of scale in purchasing, maintenance and staffing.
 - *Consistency in Care:* A single, unified system ensures that all healthcare providers access the same set of data, leading to consistent patient care regardless of location.
 - *Easier Compliance with Regulations:* Centralized systems simplify the process of adhering to legal standards and regulations, as there is a single system to audit and update.
 - *Improved Data Analysis and Research:* Centralized data storage facilitates large-scale data analysis and research, allowing more effective public health surveillance and faster identification of health trends.
 - *Enhanced Disaster Recovery:* Centralized systems can more easily implement comprehensive disaster recovery and data backup solutions.
 - *Integrated Approach to Healthcare Delivery:* Centralized EHRs are better suited to integrate various aspects of healthcare delivery, including telemedicine, patient portals and electronic prescriptions, providing a more holistic approach to patient care.

These challenges underscore the need for a more central approach to EHR management, which this research aims to explore through the proposed Centralized EHR Model for preserving Privacy and Security (CEMPS). The CEMPS framework is designed to address these specific limitations of decentralized models, offering a more unified, secure, and efficient system for managing and protecting EHRs. The framework introduces security policies and techniques to regulate authorization, aligned with the operational scenarios of the existing MHR system, involving physicians, nurses, researchers, and other stakeholders. The framework introduces security policies and techniques to regulate authorization, aligned with the operational scenarios of the existing MHR system, involving physicians, nurses, researchers, and other stakeholders. To achieve this, the study thoroughly classifies different types of health information, including personal and sensitive data, and subject them to scrutiny against both local and global privacy standards, such as HIPAA and the EU General Data GDPR. This framework facilitates the seamless sharing and use of health information between multiple stakeholders, striking a delicate balance between unfettered access and robust privacy and security. The implementation of this framework will undergo a comprehensive quantitative and qualitative

analysis, further validated through a series of illuminating case studies. Finally, in this proposed system, users exercise complete control over their health information through tailored privacy settings. In a world where healthcare data are the linchpin of patient care and the privacy of individuals is a paramount concern, the development of the CEMPS represents a critical step towards preserving patient privacy and, by extension, advancing the healthcare sector into an era of secure and efficient data sharing.

1.5.2 Justification for Centralized EHR Systems

The choice between centralized and decentralized EHR systems is pivotal in shaping the healthcare data ecosystem. Centralized EHR systems, characterized by a unified data repository managed by a single entity, offer distinct advantages in terms of interoperability, data integrity, and governance. This section elucidates the rationale for preferring centralized systems, underpinned by literature and empirical observations.

Enhanced Interoperability is a hallmark of centralized EHR systems. Unlike decentralized models where disparate systems may hinder seamless data exchange, centralized EHRs facilitate unified access to patient records across different healthcare providers. This interoperability is critical to ensure comprehensive care coordination, particularly for patients with complex or chronic conditions [92, 93].

Improved Data Integrity and Quality . Centralized systems allow a consistent application of data standards and quality controls, ensuring that health records are accurate, complete, and up to date. This is instrumental in minimizing errors and improving patient safety. Moreover, centralized architectures facilitate easier implementation of updates and security measures throughout the system, thus maintaining high data integrity standards [94, 95].

Streamlined Governance and Compliance with healthcare regulations such as HIPAA and GDPR are more straightforward in centralized EHR systems. Centralized governance enables uniform policy enforcement, privacy protections, and compliance monitoring, reducing the complexity and cost associated with managing these aspects in multiple decentralized systems [96, 97].

Real-World Efficacy . Studies and implementations in various healthcare settings have demonstrated the practical benefits of centralized EHR systems. For example, a centralized EHR initiative in a multihospital network led to significant improvements in

emergency response times and patient outcomes by providing instant access to critical patient information [95, 98]. Another study highlighted cost savings through the reduction of redundant tests and streamlined administrative processes in a centralized system [95, 97, 98].

While decentralized models offer advantages in terms of resilience and patient-centric control, the overarching benefits of centralized EHR systems in improving interoperability, ensuring data quality, simplifying governance, and demonstrating real-world efficacy provide a compelling justification for their adoption. The ongoing evolution of healthcare technology and policy will require a continuous evaluation of the design and implementation strategies of these systems to maximize their benefits for healthcare care delivery.

1.6 Research Motivation

The motivation behind this research is the rapidly evolving landscape of healthcare technology, where EHR has become a cornerstone. Despite their widespread adoption, significant concerns remain about the security and privacy of these systems. In the face of increasing cyber threats and increased data breaches, the need for a robust, secure, and private EHR system is more pressing than ever [8, 36, 87, 99–102].

Current advances in EHR technology have led to more digitized and interconnected systems, with opportunities and challenges [10–12, 14, 32, 103–106]. A detailed analysis of privacy and security concerns, including specific data breach incidents and vulnerabilities studies, illustrates the complexity of protecting patient data [13, 15–20, 52, 75, 84–86, 107–112].

A thorough review of the literature on centralized EHR systems reveals both benefits and drawbacks, particularly in terms of security and privacy [34, 35, 37, 38]. The integration of technologies like Blockchain, Cloud Computing, and Cryptography has been critical in improving the security of EHR. Recent research demonstrates their implementation and effectiveness in current systems [21, 26–31, 74, 76].

Comparative analysis with decentralized models shows that centralized EHR systems may offer superior solutions in terms of security and privacy. This comparison is supported by recent studies or data [39–41, 44, 77–79]. Regulatory and ethical considerations, such as the impact of regulations such as HIPAA and GDPR, play a crucial role in the design and implementation of secure EHR systems [46, 71].

Finally, the potential for future research is vast, including the development of more robust security protocols, exploring the use of AI to detect and prevent breaches, and studying

the long-term impacts of centralized EHR systems on patient care and data privacy. Addressing these issues is critical to building a resilient healthcare infrastructure that can withstand the challenges of the digital age and continue to provide high-quality care.

1.7 Problem Statement

Integration and optimization of EHR systems represent a turning point in healthcare technology advancement. This Ph.D. thesis specifically targets the nuanced challenges associated with the implementation of centralized EHR systems, focusing on their role in enhancing the management, security, and privacy of health records within the digital healthcare landscape.

Figure 1.1 provides a detailed visual guide mapping the intricate links between the thesis foundational problem statement, the encompassing research questions (RQ1-RQ5), and the substantial contributions derived from this study. This schematic arrangement elucidates the strategic alignment and interdependency among these critical components, emphasizing the cohesive structure that underpins the academic rigor of research.

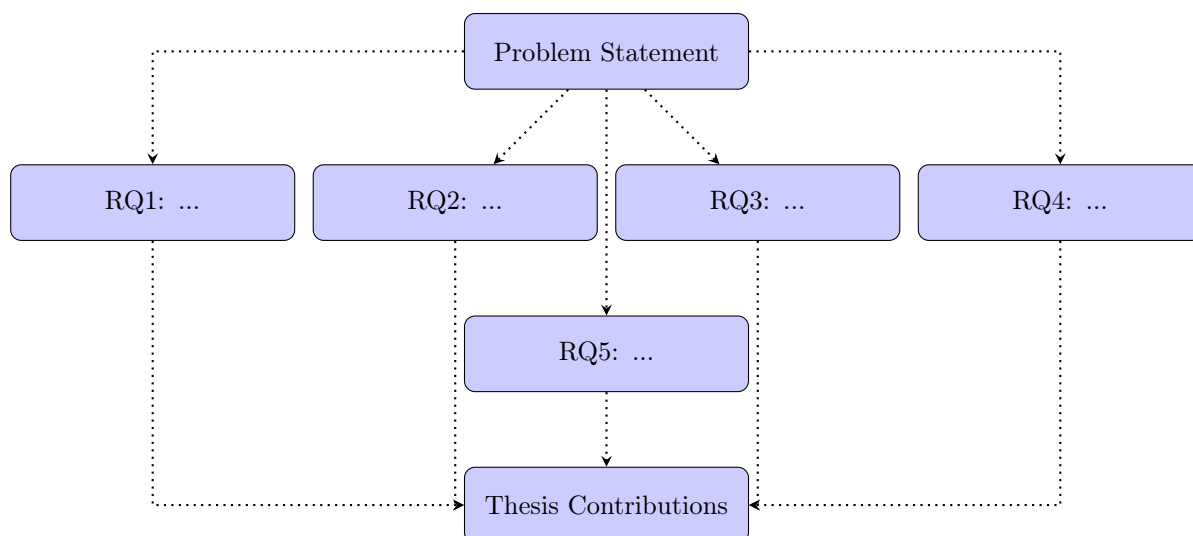


FIGURE 1.1: Interconnection between Problem Statement, Research Questions (RQ1-RQ5), and Thesis Contributions

Recognizing EHRs as a cornerstone of contemporary healthcare delivery, there is a pronounced need to analyze how a centralized framework could potentially elevate the protection and handling of patient data. This necessity is acutely felt in the face of incorporating advanced technological constructs, such as access control, blockchain, cloud computing, and cryptography, within EHR systems, as explored in [Chapter 3](#) and [Chapter 4](#).

The proposition that centralized EHR systems can offer a unified and secure data management infrastructure, thereby reducing the vulnerabilities associated with dispersed systems, sets the foundation for this thesis. However, elucidating the full spectrum of advantages and operational dynamics of centralized EHR systems, particularly how they leverage technological advances for data security and privacy, remains a critical endeavor.

This thesis aims to demystify the potential and limitations of centralized EHR systems without extensively juxtaposing them against decentralized models. It seeks to provide a detailed understanding of the architectural, security, and privacy considerations of centralized EHR frameworks. This exploration encompasses [Chapter 2](#)'s insights on EHR privacy and security challenges, [Chapter 4, Part A](#) and [Part B](#)'s discussion on the integration of innovative technologies, and culminates in [Chapter 5](#)'s examination of a universal privacy model through ontology and machine learning approaches.

Through this focused investigation, outlined comprehensively in [Chapter 6](#), the thesis aims to contribute to the development and implementation strategies of centralized EHR systems. It aims to fortify the protection of sensitive health information against the backdrop of an evolving digital healthcare framework, presenting a model that underscores the efficacy and necessity of a centralized approach to the management of EHRs in the current technological epoch.

1.8 Research Questions

Keeping in mind the preservation of privacy in the EHR and the research objectives, we formulate the following research questions to identify and review the current state of research on the preservation of privacy in the EHR.

- Q1: Can the implementation of access control, blockchain, cloud and cryptography (ABC) technologies enhance the effectiveness of EHR data sharing and access in the modern healthcare landscape? Chapters [\[2,3, 4, 5\]](#).
- Q2: Is the importance of privacy considerations in EHR data sharing among diverse stakeholders directly correlated with the distributed or centralized nature of data, leading to measurable improvements in the ethical and legal foundations of patient data management? Chapters [\[2,3, 4\]](#).
- Q3: Do the fundamental attributes of EHRs (Comprehensiveness, Accessibility, and Integration) directly correlate with measurable advantages and limitations, providing a comprehensive understanding of their role in modern healthcare? Chapters [\[1, 2, 5\]](#).

- Q4: Can the distinct contributions of EHR privacy, confidentiality, and security be quantifiable linked to the safeguarding of patient information, establishing a robust foundation for ethical and secure healthcare data management? Chapters [2,3, 4, 5].
- Q5: Can the privacy and security of EHRs be optimally ensured through the strategic integration of (ABC), addressing challenges associated with data breaches and unauthorized access? Chapters [2,3, 4, 5].

This thesis comprehensively addresses the research questions posed, demonstrating a thorough investigation and analysis of key aspects in the field of EHR systems. Specifically, the strategic integration and effectiveness of Access control, Blockchain, Cloud, and Cryptography technologies (Q1) in improving EHR data sharing and access are thoroughly examined in Chapters [2], [3], and [4], and partially in Chapter 1]. Critical analysis of privacy considerations and their correlation with the distributed or centralized nature of EHR data (Q2) is systematically explored in Chapters [2], [3], [4], and with partial coverage in Chapter1]. The correlation between the fundamental attributes of EHRs and their measurable impacts in modern healthcare (Q3) is described in the chapters [2], [3], and [4], with partial coverage in the chapters [1] and [5]. The quantifiable links between EHR privacy, confidentiality, and security in protecting patient information (Q4) are discussed in Chapters [2], [3], [4], and [5], with partial coverage in Chapter [1]. Lastly, the potential of optimally ensuring the privacy and security of EHRs through the strategic integration of (ABC) technologies (Q5) is thoroughly investigated in Chapters [2], [3], [4], and [5], with partial coverage in Chapter1]. Each of these questions is addressed with a high level of academic rigor, contributing significantly to the field and paving the way for future research directions.

1.9 Research Aim

The primary goal of this Ph.D. thesis is to enhance our understanding and application of Electronic Health Records (EHRs) in the modern healthcare landscape, specifically through the lens of a centralized EHR system's role in bolstering the security and privacy dimensions of EHR management. Recognizing the integral role of EHRs in current medical practices, this research is committed to designing innovative, technologically advanced solutions. It focuses on the synergistic application of **Access Control**, **Blockchain**, **Cloud Computing**, and **Crystalgraphy (ABC)** technologies to improve the efficacy and security of EHR data sharing and access.

A significant portion of this investigation, as detailed in [Chapter 3](#), "A SECURITY AND PRIVACY COMPLAINT DATA SHARING SOLUTION FOR HEALTHCARE DATA ECOSYSTEMS: CEMPS (CENTRALIZED EHR MODEL TO PRESERVE PRIVACY AND SECURITY)," delves into the advantages and challenges associated with centralized EHR systems versus decentralized counterparts. This analysis is crucial for advocating for a secure and privacy-focused healthcare data ecosystem built around a centralized framework.

Expanding the research's scope, [Chapter 4](#), "INTEGRATING ADVANCED TECHNOLOGIES AND ONTOLOGY MODELS FOR ENHANCED SECURITY IN ELECTRONIC HEALTH RECORDS," is divided into two enriching parts that significantly contribute to the thesis's aims:

- [Part A](#), "GPT, Ontology, and CAABAC: A Tripartite Personalized Access Control Model Anchored by Compliance, Context, and Attribute" explores the integration of cutting-edge technologies to refine access control within EHR systems. This part underlines the research's commitment to developing a nuanced, technology-driven approach to EHR privacy and security, showcasing the novel application of GPTs, ontology, and CAABAC in creating a personalized, secure access framework.
- [Part B](#), "Enhancing Health Information Systems Security: An Ontology Model Approach," further extends the thesis's exploration into security models. It presents an innovative ontology-based framework aimed at addressing gaps in current EHR security measures, highlighting the thesis' contribution to advancing health information system security through ontology and role-based access control models.

This comprehensive approach not only tackles the technological facets, but also [\[3.5.4\]](#), as further explored in [Chapter 5](#), "TOWARDS A UNIVERSAL PRIVACY MODEL FOR ELECTRONIC HEALTH RECORD SYSTEMS: AN ONTOLOGY AND MACHINE LEARNING APPROACH" considers the broader implications for policy, regulation, and ethical considerations in healthcare data management. The research aims to bridge existing gaps, clarify essential concepts, and propose a strategic fusion of methods that empower stakeholders in fully utilizing centralized EHR systems.

The desired outcome is a framework that ensures the highest level of patient data protection while maximizing the benefits of a centralized system's efficiency and integration. As described in [Chapter 6](#), "CONCLUSION AND FUTURE RESEARCH," this thesis aims to pave the way for EHRs to play a crucial role in providing high-quality, secure, and privacy-compliant healthcare services, promoting a culture of trust within the healthcare sector, and improving patient health outcomes and healthcare providers' efficiency.

1.10 Research Objectives

The objectives outlined below aim to systematically address and measure progress against identified research gaps, employing quantifiable metrics where possible. These objectives are essential to advance the understanding and implementation of secure and privacy-compliant EHR systems in the healthcare domain.

- **Centralized EHR Model Development:** Investigate and propose a robust centralized model for efficient sharing and accessing of EHR data, leveraging the latest technological advancements. This includes the exploration of Access Control, Blockchain, Cloud Computing, and Cryptography as foundational elements, as elaborated in [Chapter 3](#), [Chapter 4](#), [Part A](#) and [Part B](#).
- **Privacy in EHR Data Sharing:** Examine the critical importance of privacy considerations within the intricate network of EHR data sharing, particularly among various stakeholders. Assess how these considerations are instrumental in protecting patient information, with insights drawn from [Chapter 2](#) and further analysis in [Chapter 5](#).
- **Understanding EHRs:** Conduct an in-depth analysis of the core attributes, benefits, and challenges associated with EHR systems. The aim is to offer a holistic view of their implications for the healthcare industry, integrating findings from [Chapter 2](#) and [Chapter 3](#) to obtain a complete understanding.
- **Privacy, Confidentiality, and Security in EHRs:** Distinguish and deep investigate the unique contributions of privacy, confidentiality, and security measures in EHR systems to the protection of patient data. This involves a meticulous examination of their roles and interplay within healthcare systems, as discussed in [Chapter 4](#), [Part A](#), [Part B](#), and [Chapter 5](#).
- **Innovative Technologies for EHR Security:** Identify and evaluate cutting-edge technologies and novel approaches that are at the forefront of improving EHR privacy and security. Focus on the practical application and efficacy of these solutions in real world settings, specifically in the advancements proposed in [Chapter 4](#), [Part A](#), [Part B](#), and the universal privacy model explored in [Chapter 5](#).

These objectives align with the research questions posed at the beginning of this thesis, collectively driving toward the primary objective of contributing significant new knowledge in the realms of EHR security and privacy. Each objective is designed not only to bridge existing research gaps, but also to pave the way for future innovations in healthcare information management.

1.10.1 Comparative Analysis and Evaluation Chapter 5

Chapter 5 conducts a rigorous comparative analysis of the CEMPS framework against existing EHR privacy and security solutions. This analysis highlights the superiority of CEMPS in several key areas, including scalability, adaptability to regulatory changes, and the facilitation of secure data sharing among diverse healthcare stakeholders. The evaluation, grounded in empirical data and robust statistical methods, unequivocally demonstrates the advantages of the CEMPS framework in enhancing the privacy and security of EHR systems.

In summary, this thesis contributes significantly to the field of healthcare data privacy and security by providing a nuanced understanding of existing privacy preservation methods, introducing the innovative CEMPS framework, and empirically validating its superiority over existing solutions. These contributions mark a significant step forward in addressing the complex challenges of EHR privacy and security.

1.11 Contribution of the Thesis

This thesis offers a detailed exploration of EHR technologies, with a focus on developing a theoretical and practical understanding that drives a motivational and analytical framework [107–110]. Through a rigorous literature review in [Chapter 2](#) and subsequent chapters, this work synthesizes existing research, uncovering insights into the architectures and computational strategies of EHR systems. It critically evaluates the role of Blockchain technology in securing EHR data sharing processes, addressing the complexities surrounding data ownership and access management [111, 112].

Key contributions include:

- Systematic identification of existing gaps in the literature concerning the simultaneous achievement of data sharing and privacy preservation, initiating the discourse in [Chapter 1](#) and [Chapter 2](#).
- Detailed examination of the constructs *Privacy, Confidentiality and Security* (PCS), exploring their distinctions and overlaps, and situating this analysis within [Chapter 4, Part A](#).
- An extensive analysis of the security challenges inherent in EHR systems, proposing robust solutions to protect patient data integrity and enhance system dependability, covered in detail in [Chapter 4, Part B](#), and [Chapter 5](#).

- A critique and synthesis of various methodologies, frameworks, and technologies for EHR data sharing, advocating for an integrative approach that leverages access control mechanisms, blockchain technology, cloud services, and encryption methods, as discussed in [Chapter 2](#), [Chapter 3](#) and [Chapter 4, Part A](#) and [Part B](#).
- Development and evaluation of a new paradigm of EHR data sharing, specifically designed for interoperability with the Australian Medical Health Record (MyHR) system, to promote effective information exchange among healthcare stakeholders, developed in [Chapter 4, Part A](#) and [Chapter 5](#).
- In-depth analysis of the proposed CEMPS model for EHR management, examining its feasibility, scalability, and security aspects, thoroughly investigated in [Chapter 3](#), [Chapter 4](#) and [Chapter 6](#).
- An exploration of underresearched areas within the domain of health and medical data sharing, proposing avenues for future research that promise to advance the field, as highlighted in [Chapter 6](#).

1.12 Novelty and Superiority of the Research

This thesis delineates several groundbreaking contributions to the realm of Electronic Health Records (EHR) security and privacy, with a particular focus on [Chapters 2](#), [3](#), [4](#), and [5](#). The novelty of the research lies in the systematic survey of privacy preservation methods, the development and evaluation of the Centralized EHR Model for Preserving Privacy and Security (CEMPS), and the comparative analysis of this model against existing frameworks.

1.12.1 Systematic Survey of Privacy Preservation Methods [Chapter 2](#)

[Chapter 2](#) presents a comprehensive survey that systematically categorizes and evaluates existing privacy preservation methods in the management of healthcare data. Unlike previous surveys, this work employs a novel categorization framework that considers not only the technological aspects but also regulatory compliance and practical applicability in healthcare settings. This approach provides a multidimensional understanding of the strengths and weaknesses of current methods, laying a solid foundation for the development of the CEMPS framework.

1.12.2 CEMPS Framework: A Paradigm Shift in EHR Privacy and Security Chapters 3 and 4

The core of this thesis, detailed in Chapters 3 and 4, introduces the CEMPS framework, an innovative solution designed to enhance the security and privacy of EHR systems. CEMPS stands out for its integration of Federated Learning (FL) and Differential Privacy (DP) to ensure data privacy while maintaining the utility of healthcare data. This integration is not only novel, but also superior in terms of providing a balanced approach to privacy preservation and data utility, addressing the limitations of existing models that often compromise one for the other.

1.13 Thesis Outline

Spanning six chapters, this thesis systematically explores the realms of privacy and security within EHR, presenting innovative solutions and frameworks to enhance data confidentiality and system integrity in healthcare information systems.

[Chapter 1](#), titled "INTRODUCTION," sets the stage by defining the research's motivation, problem statement, questions, aims, and objectives. It sets out the foundational contributions of the study and provides a structured overview of the thesis content, guiding the reader through the subsequent chapters.

[Chapter 2](#), "PRIVACY PRESERVATION OF ELECTRONIC HEALTH RECORDS IN THE MODERN ERA: A SYSTEMATIC SURVEY", conducts an exhaustive review of current practices and methodologies in EHR privacy and security. This chapter evaluates the effectiveness of existing data sharing methods, assesses the roles of key stakeholders, and identifies the strengths and limitations of current EHR technologies.

[Chapter 3](#), "A SECURITY AND PRIVACY COMPLAINT DATA SHARING SOLUTION FOR HEALTHCARE DATA ECOSYSTEMS: CEMPS (CENTRALIZED EHR MODEL TO PRESERVE PRIVACY AND SECURITY) proposes CEMPS as a novel framework for secure and private data sharing within healthcare ecosystems. This chapter examines the model's security and privacy preservation mechanisms and discusses its potential benefits and drawbacks for EHR systems.

[Chapter 4](#), "INTEGRATING ADVANCED TECHNOLOGIES AND ONTOLOGY MODELS FOR ENHANCED SECURITY IN ELECTRONIC HEALTH RECORDS," is divided into two parts:

- [Part A](#), "GPT, ONTOLOGY, AND CAABAC: A TRIPARTITE PERSONALIZED ACCESS CONTROL MODEL ANCHORED BY COMPLIANCE, CONTEXT, AND ATTRIBUTE" examines the integration of Generative Pretrained Transformers, ontology, and Contextual Attribute-Based Access Control to refine access decisions within EHR systems. This part delves into the theoretical underpinnings and practical implications of the proposed framework, highlighting compliance, context sensitivity, and attribute specificity.
- [Part B](#), "ENHANCING HEALTH INFORMATION SYSTEMS SECURITY: AN ONTOLOGY MODEL APPROACH," extends the discussion to ontology models' role in bolstering EHR security. It explores conceptualization, application potential, and how it complements existing role-based access control systems.

[Chapter 5](#), "TOWARDS A UNIVERSAL PRIVACY MODEL FOR ELECTRONIC HEALTH RECORD SYSTEMS: AN ONTOLOGY AND MACHINE LEARNING APPROACH," outlines the development of a universal privacy model leveraging ontology and machine learning. This chapter argues for a comprehensive and adaptive approach to EHR privacy, supported by advanced technological solutions.

[Chapter 6](#), "CONCLUSION AND FUTURE RESEARCH", synthesizes the research findings, reiterates the study's contributions, and discusses its limitations. Critically examines the implications of the research for future studies in EHR security and privacy, suggesting directions for forthcoming investigations.

This thesis aims to contribute significantly to the privacy and security enhancements in EHR systems, offering a deep, multifaceted analysis of critical issues and proposing forward-thinking solutions.

Chapter 2

PRIVACY PRESERVATION OF ELECTRONIC HEALTH RECORDS IN THE MODERN ERA: A SYSTEMATIC SURVEY

NOTE: The content of this chapter has been published in *ACM Computing Surveys*. Nowrozy, R., et al. (2023, July). Preservation of Privacy of Electronic Health Records in the Modern Era: A Systematic Survey, *ACM Computing Surveys*. <https://dl.acm.org/doi/abs/10.1145/3653297>. Incorporated as Chapter 2.

2.1 Introduction

Electronic Health Records (EHR) contain data on a patient's medical history in digital form; therefore, it is extremely important that EHRs are secure, with privacy and confidentiality being its key goals. The literature indicates the many benefits of EHRs over traditional paper-based records, such as cost savings, improved quality of healthcare care, advancement of evidence-based medicine, data collection, and flexibility [7, 8, 14, 32, 87, 99, 101, 102, 113–117]. To realize these benefits, EHR systems must satisfy certain requirements and follow several criteria with respect to data completeness, resilience to failure, high availability, and consistency of security policies [103].

A crucial aspect in the development and implementation of EHR systems is the understanding and distinct treatment of Privacy, Confidentiality, and Security (PCS). Historically, these terms have been used interchangeably in the literature, yet their differences have significant implications for EHR solutions [7, 114, 118, 119]. Our review of 130 studies on EHR and privacy preservation techniques, spanning 2012 to 2022, aimed to clarify these distinctions and explore the methods employed to preserve the privacy of EHRs [8, 102, 115].

Studies describe EHR systems as unreliable, resulting in compromised patient privacy [113, 120, 121]. In the past, data and privacy breaches in the healthcare sector exposed 112 million records [105]. A survey in 2018 found health breaches to be the most costly, surpassing other sectors [8]. Privacy and security acts emphasize the importance of EHR privacy [7, 10, 102, 104, 105, 122, 123].

In response to these challenges, our findings suggest that access control, blockchain, cloud-based, and cryptography techniques are commonly employed for EHR data sharing [32, 101, 102]. We have also summarized commonly used strategies and collated a comprehensive list of differences and similarities between PCS [7, 10]. Furthermore, we propose a fusion of techniques to enhance PCS in EHRs, summarized in a tabular form for clarity and ease of understanding [116, 117].

This chapter presents a systematic review of the literature on the tools and methodologies used to protect the privacy of EHR. The literature review was carried out using the Kitchenham methodology [124] and examines around 130 publications from 2012 to 2022. This chapter contributes to the literature on the privacy preservation of EHRs for researchers interested in this domain. First, it shows how EHR privacy can be breached. Second, it helps to understand the concepts of *privacy*, *confidentiality*, and *security* (PCS) as separate terms. Third, it presents how different technologies can be implemented to ensure data privacy. Fourth, it identifies areas for future study that require greater attention so that practitioners and researchers can generate ideas to improve the privacy of the EHR. However, more research is needed to clarify the understanding of PCS in relation to EHR.

The rest of this survey is structured as follows. The motivation and background of the survey is explained in section 2.2. The information in Section 2.3 describes the survey plan and conduct. Section 2.4 covers the reporting and analyzing of the results. Section 2.5 contains the survey findings, discussion, and limitations. Section 2.6 concludes the chapter.

2.2 Background

In this section, we introduce the conceptual terms used in the literature review, namely EHRs and preservation of the privacy of EHRs. Finally, reviews of the literature on different technologies and the preservation of EHR privacy are briefly reported to explain the motivation behind this research. **Table 2.1** lists the key words and the glossary used in this chapter.

TABLE 2.1: PCS: Privacy, Confidentiality, and Security.

| Glossary | Description |
|-------------------------------------|---|
| Privacy | In the context of EHR, privacy is defined as the right of individuals to keep their health information confidential and to control the access and use of these data. It is a fundamental right of the patient under various health laws and regulations. |
| Confidentiality | Refers to the ethical and legal duty of healthcare professionals to protect personal health information from unauthorized disclosure. It is critical to maintain trust between patients and healthcare providers. |
| Security | Encompasses the technical and organizational measures to protect EHR data from unauthorized access, use, disclosure, disruption, modification, or destruction. Security practices are crucial to maintaining the integrity and availability of health data. |
| EHR Management | Involves the systematic approach to the management and governance of EHR systems, focusing on efficient and secure data handling, storage, and exchange, ensuring compliance with legal and ethical standards. |
| EHR Systems and Technologies | Refers to the hardware, software and methodologies used in EHR systems. This includes traditional and emerging technologies such as cloud computing, blockchain, and AI-driven analytics used to improve EHR functionality. |
| PCS Framework | Represents the integrated approach to Privacy, Confidentiality, and Security in EHR systems. It underscores the interrelatedness of these aspects in ensuring the holistic protection and governance of health information. |

2.2.1 EHR

EHRs comprise patient data in digital form that are securely stored and exchanged and accessible by multiple authorized users to support the continuous and efficient management of integrated healthcare [12–14]. EHRs comprise details of patient medical histories, including diagnosis, laboratory findings, information about hospital admissions, surgical procedures, and medications. They describe the patient’s condition, allowing for a more detailed diagnosis and treatment of the patient [15]. EHRs can be shared with other healthcare providers when necessary. However, EHRs are prone to various types of security and privacy attacks during transmission [15, 125]. In light of its widespread use, the development of a safe environment for the sharing of EHR has gained a lot of interest in the healthcare sector. The most recent literature [126–128] indicates that there are many benefits to using EHR software, including cost savings, increased quality of healthcare care, advancement of evidence-based medicine, more comprehensive data collection, and flexibility. Consequently, the term EHR in this chapter refers not only to an electronic database to store and retrieve health information, but also to a system that can be used to enforce and maintain completeness of data, resilience to failure, high availability, and

consistency of security policies. Finally, we acknowledge the different nature of health information such as *Personal Health Records* (PHRs) and *Electronic Medical Records* (EMRs) and their privacy, but, in this chapter, we focus specifically on the EHR.

2.2.2 EHR Privacy

EHR privacy refers to the protection of patients' rights over their data, encompassing both data protection and physical privacy. It involves ensuring that patients have control over their health-related data, maintained under stringent privacy and security policies [123]. Privacy in EHRs also includes mechanisms to track data access and transmission, protect against social or economic discrimination, and foster trust in healthcare systems.

The privacy rights of a patient encompass both their data and their physical privacy. Trust between healthcare workers and patients is fundamental to the practice of medicine. The patient must trust the doctor enough to share personal details that can be stressful, embarrassing, or potentially damaging. A physician must trust that a patient is sharing enough information to make an accurate diagnosis and that a patient can give informed consent to treatments that may pose significant risks [129]. An essential component of trust between the doctor and the patient is privacy. More than two thousand years ago, Hippocrates emphasized the importance of privacy, and the practice of medicine has recognized and valued the importance of privacy ever since [130, 131]. Privacy is one of the main cybersecurity challenges today, and privacy concerns are regularly addressed in the ubiquitous healthcare system by researchers and end users. Patients who use EHR systems should have control over their health-related data and these data should be maintained by stringent EHR privacy and security policies at national and global levels. This must include actions for compensation for data breaches that have occurred, not just for those at risk. By doing this, protection from social or economic discrimination and building trust in the health care system can be achieved. However, it is necessary to ensure that critical health data remain accessible at the point of care and that systems are in place to manage privacy protection. Control of patients' own data requires appropriate privacy-preserving systems that can also help track who has viewed a record and to whom it has been transmitted [123]. Regarding information privacy, a patient has the right to know the personal health data collected on him and the way it has been used.

Software systems handling personal and important user data such as EHRs are facing difficulties in ensuring a high level of data privacy [12]. Health-related information should only be accessed or used by authorized and approved users, such as medical practitioners, as it is confidential and sensitive data. To ensure the safety and protection of user

data, extensive rules and standards have been proposed. Strict security measures are in place to govern the transmission of health data which will result in severe penalties for non-compliance [12]. In many countries around the world, health information is centralized at a national level, for example, with the National Health Scheme in the UK. Whenever a *General Physician* (GP) updates patient registration information on their clinical system, *Primary Care Support England* (PCSE) uses this information to update the *National Health Application and Infrastructure Services* (NHAIS) which holds the National Patient Register. However, NHAIS is different from GP clinical systems and PCSE cannot see data on GP systems due to privacy concerns [131]. The *Royal Australian College of General Practitioners* (RACGP) also created a sample registration form for new patients. To ensure the privacy of health information, as required by federal and state privacy laws, the form complies with the RACGP standards for general practices (5th edition) [130]. If patients have privacy concerns, they can discuss them with their GP and leave the form blank. But it is not considered a good practice to let the patient leave the form blank, as the information may be required at any phase of their treatment, and missing data may result in incorrect treatment.

2.2.3 EHR Confidentiality

Confidentiality in EHRs is related to the protection of identifiable personal health information, shared only with explicit informed consent. It ensures that sensitive data are protected from unauthorized disclosure [132]. Measures such as data encryption and the adherence to privacy laws are key to maintaining confidentiality in EHR systems [54]. Confidentiality involves the protection of recognizable personal information. It is an agreement and informed consent procedure that guarantees that an individual's identity and personal data will only be shared with another individual or department with their express informed consent [132–135]. The fact that data confidentiality is unavoidable should be made known. The access to data can put one's confidentiality at risk, therefore to ensure the confidentiality of electronic data, it must be encrypted [132]. Confidentiality is one of the core concepts of cybersecurity and ensures that private information is protected from unauthorized disclosure [54].

2.2.4 EHR Security

EHR security focuses on protecting health information from unauthorized access, misuse, or breaches. Covers authentication, authorization, and access control measures [6, 59]. Security in EHRs involves technical and administrative strategies, with compliance with

standards such as HIPAA and HITECH being crucial [136]. Security ensures the integrity, confidentiality, and availability of health information.

EHRs are shared among different systems, raising concerns about patient privacy due to the possibility of unauthorized access or misuse due to improper security implementation including authentication, authorization, and access control [6, 59, 137]. Defining access control strategies and policies is crucial to secure EHR systems. Data security involves the protection of personal information against accidental or unlawful destruction or accidental loss, alteration, unauthorized disclosure, or access. Ensuring the security of EHR systems has been an important aspect in designing, implementing, and managing the shared care paradigm; the requirements for such security and privacy of EHRs need to be identified to be applicable in EHR systems. To ensure the security and privacy of EHR while providing shared and interoperable EHR services, healthcare organizations have highlighted the importance of standards [64, 138]. Examples of such standard developers and publishers include: *Health Level Seven (HL7)*, *HIPAA* and *Health Information Technology for Economic and Clinical Health Act (HITECH)* in the USA; *Canada Health Infoway* in Canada; *HEASNET* in Japan; and *ISO/TC 215, CEN/TC* in Europe [136]. *ISO 27799* focuses specifically on the information security management perspective for EHR security rather than the technical perspective. EHR requires interoperability, which requires information security, including the restriction of unauthorized access, use, disclosure, and modification of data to ensure confidentiality, integrity, and availability [139]. EHR connects through wireless communication protocols which can generate a massive amount of data at regular intervals, opening the doors for attackers to launch various security attacks. An insecure technique for Healthcare 4.0 [140] may lead to a breach of healthcare records where hackers can gain full access to patient email accounts, messages and reports [141]. Security procedures are used to control access to patient data to protect it from unauthorized users. This can be achieved with operational controls within a privacy-protected entity [137, 142].

2.2.5 The Distinction Between EHR Privacy, Confidentiality and Security

Understanding the distinctions between EHR privacy, confidentiality, and security is crucial for comprehensive protection of patient data. These concepts are intimately related, yet distinct in their application and point of compromise in the event of a cyber attack, as described by the Cyber Kill Chain (CKC) model¹.

¹<https://www.lockheedmartin.com/en-us/capabilities/cyber/cyber-kill-chain.html>

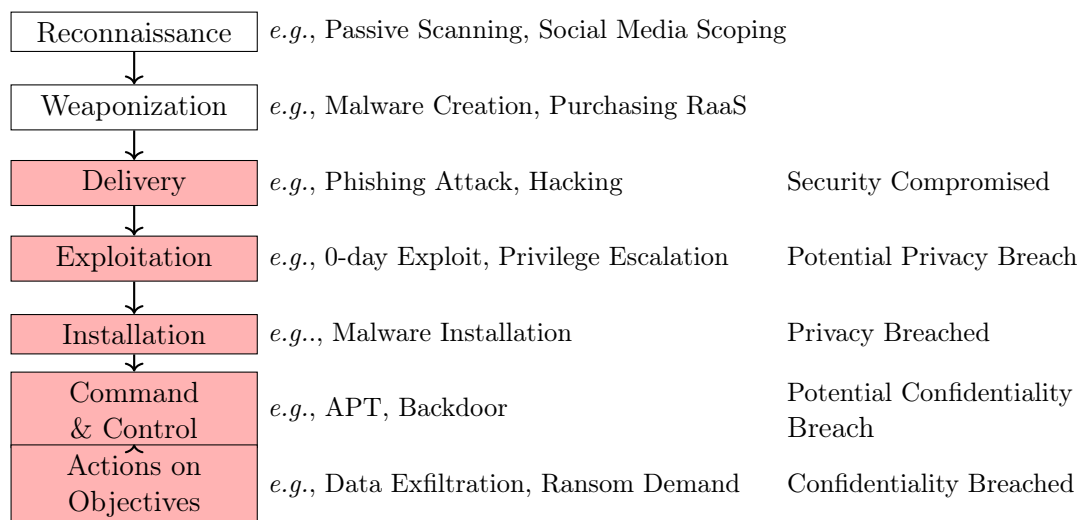


FIGURE 2.1: The Cyber Kill Chain illustrating the points of compromise for Security, Privacy, and Confidentiality in EHR systems

- EHR Privacy:** This refers to the patient’s ability to control their health information and protect it from unauthorized access. Privacy is compromised when attackers gain unauthorized access to EHR systems, which can occur at the CKC’s Installation phase. However, privacy is directly violated when attackers have the ability to read private patient information, which may happen during *Exploitation*, but privacy breach is fully realized in the *Installation* phase if the data is accessed.
- EHR Confidentiality:** Refers to the ethical and legal obligation to keep health information accessible only with the patient’s explicit consent. A potential breach of confidentiality occurs during the *Command & Control* phase when attackers have the ability to exfiltrate sensitive data. However, confidentiality is definitively compromised in the *Actions on Objectives* phase if the data is actually extracted from the EHR system.
- EHR Security:** This involves the technical and administrative safeguards that protect EHR systems from unauthorized access, data breaches, and cyber threats. Security is first compromised in the *Delivery* phase of the CKC when the attack vector is successfully deployed into the healthcare system.

This alignment with the CKC phases (Figure 2.1) illustrates that while security can be compromised by the mere success of a delivery mechanism, privacy is specifically breached when unauthorized viewing of patient information occurs, and confidentiality is breached when information is extracted and possibly used for malicious intent such as blackmail or public disclosure.

2.3 The Survey Plan and Conduct

We used a systematic, comprehensive, reproducible strategy to review articles on methods and technologies related to EHR privacy to identify and categorize them. We carry out the survey in three stages, *that is*, plan, conduct, report [124]. This section discusses the planning of the survey and how it was conducted.

2.3.1 Planning the Survey

Planning included the following activities:

2.3.1.1 Constructing the Survey Questions

The survey aims to explore the current state of privacy preservation in EHR systems. The breakdown into five specific survey questions (SQs) serves to provide a comprehensive understanding by covering various aspects: the methods of sharing EHR data, the role of privacy in stakeholder engagement, the strengths and weaknesses of EHR systems in privacy preservation, the distinction between privacy, confidentiality, and security in EHRs, and the technologies available for maintaining EHR privacy.

SQ1: What EHR data sharing methods are currently available?

Justification: This question is intended to catalog existing EHR data sharing methods, establishing a foundational understanding of the mechanisms through which privacy must be maintained.

SQ2: What role does privacy play when sharing EHRs with different stakeholders?

Justification: By examining the role of privacy in stakeholder engagement, this question seeks to highlight privacy expectations and requirements from various perspectives within the healthcare system.

SQ3: What are the main strengths and weaknesses of EHR systems in terms of preserving privacy?

Justification: Identifying the strengths and weaknesses of current EHR systems provides information on their privacy preservation capabilities and the areas that need improvement.

SQ4: What is the difference between EHR privacy, confidentiality, and security?

Justification: Clarifying the distinctions between these concepts is crucial, as each has unique implications for the design of privacy-preserving measures within EHR systems.

SQ5: What different technologies are available to preserve the privacy of the EHR?

Justification: This question aims to explore the spectrum of technologies that can or are being implemented to protect the privacy of EHR, thus forming future research and development in this area.

Through five targeted questions, the main goal of the survey is to dive into EHR data sharing, stakeholder privacy roles, and the strengths and weaknesses of EHR systems, thus providing a comprehensive understanding of privacy concerns in EHR data sharing.

2.3.1.2 The Search Keywords

Following a preliminary analysis of some of the most widely read works on the topic of protecting the privacy of EHRs, as well as our own expertise, we selected several search keywords. First, the majority of articles in this field were retrieved using the acronym EHR. We also selected EMR [14], PHR, security, privacy, confidentiality, secrecy, sharing, access, and breach as significant terms, since they are often used in relation to EHRs. Furthermore, since this review focuses on the privacy of EHRs, the difference between various terms that are used interchangeably (, *e.g.* PCS) is also mentioned to limit the focus and explain the confusion among them. The following are the important terms used in this study.

- *Electronic Medical Record (EMR):* An EMR is an application environment that comprises clinical data repository, clinical decision support, controlled medical vocabulary, order entry, computerized provider order entry, pharmacy and clinical documentation applications that can be confidential in different ways to different stakeholders [143].
- *Personal Health Record (PHR):* A PHR is a collection of medical documentation of an individual maintained by the individual themselves or a caregiver in cases where the patient cannot do so themselves [14].
- *Privacy:* Privacy refers to an individual's control over how much, when, and under what circumstances they can share details of their physical, behavioral, or intellectual life with others, and their right to restrict other people's access to their personal information [144].
- *Security:* Security is the practice of protecting digital information from unauthorized access, corruption, or theft throughout its lifecycle.
- *Confidentiality:* Data confidentiality means protecting data against unintentional, unlawful, or unauthorized access, disclosure, or theft [144].

- *Secrecy*: Data secrecy means that data is completely unknown and untraceable by anyone other than the owner and those with whom it has been explicitly shared.
- *Sharing*: Data sharing is the ability to make a data resource available at various points.
- *Access*: Data access is the ability of a user to access or retrieve data stored within a database or other repository.
- *Breach*: A data breach exposes confidential, sensitive, or protected information to an unauthorized person. It is important to note that several spellings (such as behavior/behavior and modeling/modeling) were also used in the searches to make sure that no relevant publications were overlooked.

2.3.1.3 Data Collection

Data was collected via three sources *i.e.* digital libraries namely, Google Scholar, Elsevier Science Direct, Springer Link, ACM Digital Library, and IEEE Xplore, journals, and conference proceedings. We found that the search engines of the most well-known scientific libraries performed differently when the search string was specified. Depending on the library, multiple methods had to be used to perform the same search (*i.e.*, using different syntax). There were a variety of alternatives in each library to find content, for instance, by keywords in the title, abstract, or entire article. As the technologies for data sharing and privacy research are multidisciplinary, therefore, all searches were carried out comprehensively. The conference proceedings and articles relevant to the survey and the studies in the reference lists of selected articles were also examined.

2.3.1.4 Approach to Gathering Data

For our research, data was accumulated from a trio of primary sources, notably digital repositories including Google Scholar, Elsevier Science Direct, Springer Link, ACM Digital Library and IEEE Xplore, along with journals and symposium records. The selection of journals and conference papers was intentionally narrowed to those that contributed significantly to the domain of privacy and security of EHR. This included authoritative sources in the realms of healthcare informatics, data protection, and privacy legislation. A detailed enumeration of these essential journals and symposia is explained in Table 2.2.

The exploration revealed varied responses from the search engines of prominent scientific databases when specific search queries were entered. The need to adapt methodologies

TABLE 2.2: List of Target Journals and Conferences

| Journal/Conference | Selection Criteria |
|---|---|
| ACM Conference on Health Informatics | Leading conference in healthcare technology |
| Blockchain in Healthcare | Innovative applications of blockchain technology in healthcare |
| Digital Health and Telemedicine | Advances in digital health and telemedicine practices |
| Emerging Technologies in Healthcare | Exploration of new and emerging technologies in healthcare |
| Health Informatics and Data Analysis | Developments in health informatics and medical data analysis |
| Healthcare Policy and Management | Studies on healthcare policy, management, and regulatory compliance |
| IEEE Journal of Biomedical and Health Informatics | Notable for biomedical informatics research |
| International Conference on Medical Data Privacy | Focus on medical data privacy laws and practices |
| Journal of Healthcare Privacy and Security | Specialized in EHR privacy and security |
| Journal of Medical Internet Research | High impact factor, focus on digital health |
| Medical Data Security and Encryption | Research in data security and encryption for medical data |
| Patient Privacy and Rights | Research on patient privacy, rights, and ethical considerations |

for identical queries across different databases became apparent (*i.e.*, modifying search syntax). Each database offered a spectrum of search options, such as pinpointing keywords in titles, abstracts, or throughout the text. Given the interdisciplinary nature of data sharing and privacy in research, our search strategy was exhaustive. We also meticulously reviewed conference papers and journal articles that were relevant to our study, including those cited in the bibliographies of the selected primary articles.

2.3.2 Conducting Survey

The survey was carried out after the planning phase. Literature review and data synthesis are discussed in Sections 3.2.1 and 3.2.2 sequentially.

2.3.2.1 Executing Survey

1. Searching digital libraries: the digital libraries detailed in section 3.1.3 were searched using the search keywords (section 3.1.2).
2. Search for conferences and journals: All conferences and journals (Section 3.1.3) were searched using search keywords (Section 3.1.2).
3. Searching Backward Snowballing: To identify relevant papers, we searched for references and citations in the publications which were identified in the first two rounds.

2.3.2.2 Data Synthesis

The publications were listed in an Excel spreadsheet. The search of digital libraries yielded 513 relevant publications in four categories, namely, journal articles, conference articles, review articles and case studies. We only selected articles that had been published from 2012 onward, taking into account the more practical and accurate analysis of

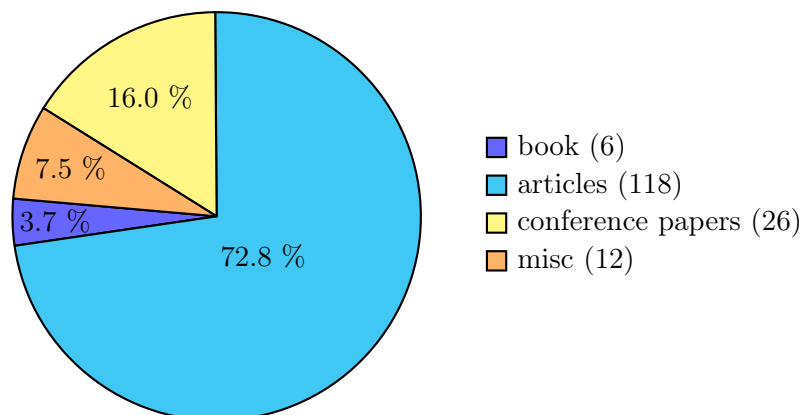


FIGURE 2.2: Numbers and Percentages of Publication Types

publications. We carefully examined each publication and categorized it into one of the four categories along with its publication year. We removed any duplicate studies that resulted in a total of 162 publications (**Figure 2.2**). Against each category, we made three further subcategories based on the main topics: a) EHR, b) EHR privacy, c) EHR security, d) EHR confidentiality, hence reducing the results to 130 publications. The selected publications for the comparative review were based on access control, blockchain, cloud-based techniques, and cryptography.

2.4 Survey Results

The analysis of the survey results is based on the survey questions (Section 3.1.1).

2.4.1 What EHR data sharing methods are currently available?

2.4.1.1 Cloud-based Sharing

Cloud-based platforms offer advantages in delivering electronic health services by providing ubiquitous network access, scalability, and cost savings [10, 145]. However, transferring EHRs to the cloud poses major threats to privacy, data integrity, and confidentiality, and additional techniques are required to ensure data secrecy. Wang and Song proposed a cloud-based EHR system that uses an attribute-based cryptosystem and blockchain technology to solve these security problems, achieving confidentiality, authentication, integrity of medical data and supporting the sharing of confidential data [145, 146]. The researchers used *Attribute-Based Encryption (ABE)* and *Identity-Based Encryption (IBE)* to encrypt data, ensuring fine-grained access control for encrypted data, and used

an *Identity-Based Signature* (IBS) to implement digital signatures. To achieve different functions of attribute and identity-based encryption and identity-based signature in one cryptosystem, they introduced a new cryptographic primitive, called combined *Attribute-Based/Identity-Based Encryption and Signature* (C-AB/IB-ES) which eliminates the need for different cryptographic systems for different security requirements. In addition, they use blockchain technology to ensure that medical data cannot be tampered with and data sources can be traced. Their technique is a well-defined and encrypted data sharing method, but its scope is limited to only patients and hospitals and does not accommodate the needs of various other health workers such as pharmacists, clinicians, researchers, etc. but this can easily be extended because of the helpful property of blockchain to trace data sources.

2.4.1.2 Attribute-based Access Control (ABAC)

ABAC enables attribute-based encryption for secure access to cloud-based EHR systems [15]. ABE ensures tight data security and records every patient visit as a separate node in the knowledge graph, facilitating easy querying and faster data access.

2.4.1.3 Role-based Access Control (RBAC)

RBAC [104] in cloud storage gives various types of users different access privileges. This policy transformation approach enables EHR data to be transferred from a private cloud to a public cloud with the corresponding transformation in the access control policy. Conditional or emergency access and authorization are delegated.

2.4.1.4 Encryption-based Sharing

Keyword searchable encryption and proxy re-encryption technology is combined in [115] for privacy-preserving and secure data sharing for EHR sharing based on consortium blockchain technology and cloud storage. Proxy re-encryption (a safe cryptographic method) is used to ensure effective access control of confidential data [147]. The re-encryption of cyphertext by the cloud is a relatively good opportunity to enhance the security of data in the technique proposed by [115]. But keyword searchable encryption is not clearly described from the user point of view. A privacy-preserving framework for the control of access and interoperability of EHRs using blockchain technology [11] is a blockchain-based framework for secure, interoperable and efficient access to medical records of patients, providers, and third parties, while preserving the privacy of sensitive

information of patients. Keyword searchable encryption and proxy re-encryption technology [115], consortium blockchain technology, cloud storage proxy re-encryption cloud technology ensures that users can find the relevant EHRs and protects data security with a searchability guarantee that only authorized entities can access the EHRs. It indicates that the challenging problem of private searching for encrypted data is of independent interest and deserves further study.

Attribute-based cryptosystems [8, 115, 148–151] encrypt data, ensuring fine-grained access control for encrypted data, and use an IBS to implement digital signatures. Introduces a new cryptographic primitive, called combined attribute-based/identity-based encryption and signature. The cloud server may not be fully trusted. A general verification mechanism that can be applied to all search schemes is also lacking in the current literature. There is also no effective countermeasure to penalize a misbehaving server or user. Security techniques for data sharing may include, but are not limited to, items such as firewalls, virus checking, encryption, and decryption, as well as authentication measures.

2.4.1.5 Blockchain-based Sharing

The blockchain-based [9, 102, 105, 147–154] data sharing mechanism [101] offers a secure distributed research data sharing network. It provides a way to specify/control the parameters of sharing and provides full accountability of access to such data. The Ancile privacy preservation framework [11] uses smart contracts in an Ethereum-based blockchain using cryptographic techniques, implementing six separate contracts, thus improving the efficiency of patient experience and reducing privacy threats. The patient can be the only node that expressly gives the location of their information. Searchable encryption based on blockchains [155] guarantees that users can receive accurate search results without additional verification. It allows cryptographic algorithms to be built to ensure data integrity, standardized auditing, and some formalized contracts for data access. Zaghlol [8] proposed a decentralized and hierarchical data sharing method using smart contracts that offers a secure, private and efficient electronic record sharing scheme that utilizes smart contracts deployed on a blockchain. It empowers patients to have control over their records, allowing them to selectively share these with data users that satisfy their privacy preferences. It also provides patients with access control over their records and eliminates the need for management services provided by record-generating parties.

2.4.1.6 HIPPA/Privacy Act-based Sharing

The HIPAA Security Rule incorporates three safeguards, namely administrative, physical and technical [156] which encompass a wide array of security techniques that are implemented by healthcare organizations to protect health information in EHRs [115, 157]. HIPPA focuses on compliance with security policies and procedures and the protection of physical access to protect health information through hardware and software access. In Europe, sharing health data and access to data is subject to *GDPR* [158] which provides subject data rights to EU citizens and is much broader in scope [102].

2.4.1.7 Current Australian MyHR (*My Health Record*) Sharing

The lack of interaction between health care practitioners and across various settings in the healthcare sector has been widely observed. Hence, it is possible that the patient's history will not be easily available at the time of admission. The implementation of My Health Record in *Emergency Departments* (EDs) was undertaken by The Australian Digital Health Agency and the Australian Commission on Safety and Quality in Health Care (ACSQHC) which will empower the use of My Health Record by medical professionals in EDs in Australia.

2.4.1.8 Model-based Sharing

To promote the sharing and integration of patient records, healthcare institutions generally follow three models: push, pull, and display [105]. Medical data is sent from one supplier to the other in a push model (, *e.g.*, from an emergency room physician to a chief care doctor). A provider asks another vendor for data in a pull model (, *for example*, a cardiothoracic specialist consults with a primary care physician). A vendor looks at the registry of another company in the display model. For example, a cardiologist reviews the X-ray of a patient that was obtained in an emergency clinic. A significant downside to these is that evidence is not inspected in a structured manner. In comparison, in the push model, a new hospital may not be able to view the information that was 'pushed' to the first hospital if a patient is moved to another hospital. The absence of an independent audit (such as the compliance audit of *Hippocratic Databases* (HDB) [6, 159]) means that the precision of the data is not ensured from the perspective of data generation to the perspective of data usage. Permission is also given on an informal and ad hoc level in the pull model. The procedures and guidelines that regulate HDB differ significantly between territories, depending, among other things, on local experience and national implementation of privacy policies [130]. This approach helps companies record

past information in a metadata format recorded in a log database. Auditors make use of these records with the help of queries to extract information such as the identity of a user who accessed a specific record, the time and date of the query, the purpose of access, and the results of the query. This HDB approach uses relevance-ranking auditing disclosures that depend on sensitive data tracking. This can lead to misuse compared to the results of previous queries stored in the backlog. Work in [160] investigated database compliance issues and improved database system accountability by saving previous data events to ensure compliance with the company's top-level policies. The adoption of such methods includes high human-based validation to perform compliance-validation tasks. This is due to the incompatibility of the proposed mechanisms to automatically determine fine-grained requirements.

2.4.2 What role does privacy play when sharing EHR with different stakeholders?

Data privacy has become an paramount concern in the realm of EHR sharing between various stakeholders. The intricacies of data privacy in distributed medical research and healthcare systems revolve around policies shaped by legislative and jurisdictional directives [16, 17, 161]. These policies require stringent enforcement at the program level, but often do not provide an ironclad guarantee of privacy protection [162, 163].

Social acceptance of healthcare systems is highly dependent on the scrutiny and improvement of privacy agreements. Implementing health mechanisms that protect privacy can substantially alleviate public concerns in Australia about the privacy of their data [20]. This requires the adoption of robust mechanisms to protect the privacy of patient data. However, despite technological advances and automation, privacy risks persist in healthcare data [6, 10].

EHRs, which are easily accessible and necessary for better patient care, pose unique challenges. Access to the EHR must be balanced with the greatest respect for privacy and confidentiality [41]. Furthermore, these systems are susceptible to cyber security risks, including threats from hackers and system failures, underscored by the need for rigorous governance and audit mechanisms in healthcare applications [148, 164].

Effective governance involves user guidelines and adherence to initial test conformance during access control run-time. This includes reviewing past data-sharing events and the permissions granted to each user. Automated solutions for such reviews, while beneficial, are often challenging and costly [159].

EHR comprises various elements such as personal, sensitive, private, and historical health information. Streamlining EHR systems for privacy assurance may benefit from standard protocols, such as the EU GDPR, which imposes stringent data protection requirements and penalties for noncompliance [165].

Patient care involves a holistic approach, respecting the preferences and values of the individual patient in clinical decisions [166]. Healthcare systems utilize sensitive patient data, which are subject to privacy laws due to their personal nature. Balancing patient data autonomy and privacy with the public benefit derived from these data is a key concern for healthcare policy makers and security developers [133].

Access to a patient's electronic health record is typically granted to various medical professionals, necessitating complex access control mechanisms. The use of pseudonymized or anonymized records for research purposes introduces challenges, especially with regard to genetic data and the necessity of fine-grained access rules [105].

Protecting sensitive data from unauthorized third-party access requires intelligent audit systems. The UK National Health System (UKNHS), for example, employs networks for auditing but faces challenges in ensuring complete and separate audits [163]. Continuous internal and external audit trials are essential to align operational processes with high-level policies and to track regulation breaches effectively.

Under GDPR, personal information is broadly defined, which includes a wide range of data that could identify an individual [11]. Sensitive data require extra safeguards due to potential risks of discrimination if not handled properly.

2.4.2.1 Differentiation of Health Record Sets.

The categorization of EHR data into distinct sets plays a crucial role in understanding the complexities of data privacy and access control. We define several Health Record Sets as follows:

- *Health Record Set A = {Personal health information}*
- *Health Record Set B = {Personal health information, Private health information}*
- *Health Record Set C = {Personal health information, Sensitive health information, Private health information, Historical health information}*
- *Health Record Set D = {Personal health information, Historical health information}*
- *Health Record Set X = {Health Record A, Health Record B, Health Record C, Health Record D}*

These categorizations are critical to understanding how various types of information intersect and combine within the healthcare system. For example, the intersection of sets A and B of health records can be represented as:

$$\text{Health Record Set A} \cap \text{Health Record Set B} = x : (x \in \text{Personal health information}) \wedge (x \in \text{Personal health information, Private health information})$$

which simplifies to the set of all personal health information. Similarly, the union of Sets B and C can be expressed as

$$\text{Set B} \cup \text{Set C} = \{x : (x \in \text{Set B}) \vee (x \in \text{Set C})\} = \{\text{PHI, SHI, PriHI, HistHI}\}$$

This mathematical representation helps visualize the overlap and unique aspects of different types of health information, highlighting the complexity of managing privacy across various data sets.

2.4.2.2 Case Studies in EHR Data Security

Professionals and respective roles Support Professionals Nursing Professionals Medical Practitioners Diagnosis Professionals Medical Scientists Receptionist, Chemist, Nurse, Nurse Manager Doctor, Specialist, Psychiatrists Radiologist, psychologists Researcher, Senior Researcher, Junior Researcher Health Record Set A Health Record Set B Health Record Set C Health Record Set D Health Record Set X. In a hybrid way, the combination of healthcare and information technology is an ongoing process, which can bring many changes to the healthcare discipline. These developments affect the recovery process of patients and therefore require diligent data collection. Healthcare is entirely based on data for service, which poses some questions regarding data access and privacy preservation. The word secrecy means allowing someone to access patient PII and also ensures that private data can only be obtained by authenticated individuals. Ensuring that these sensitive data are kept secure from eavesdroppers or trespassers is related to the term protection, which ensures that the device is capable of protecting the private data of users from strangers [113]. Therefore, the privacy risks and attack possibilities that patients' EHR data may encounter and the various techniques used to handle these attacks are discussed below.

Case I: EHR is accessed by an intruder: The authentication server will control the intruder. This authentication server uses the RADIUS (Remote Authentication Dial-In User Service) protocol, based on server/client service. The protocol is designed in such a way that users' information is passed from clients to RADIUS servers and acts based on

a returned response. In this RADIUS protocol, it first receives the connection requests of the users, followed by user authentication, and the necessary information is processed to the client to offer service delivery. Several methods are supported by the RADIUS server for user authentication. After the user logs in to this server with their user ID and password, the users are offered many authenticated mechanisms such as UNIX, CHAP, PAP or PPP login. In general, an Access Request query of the NAS and the server response (Access Reject/Accept) constitute the user log-in. The RADIUS server searches for the username from the database after receiving the NAS access request. It loads the default profile, or an Access-Reject message is sent when the RADIUS server does not find the required username from the database. The Access-Reject message is simply a text message that provides the reason for refusal.

Case II: X tries to masquerade as Y: The permissions list is maintained by the ACL server that combines the EHR data of patients such as drug-related data, neonatal data, sexual health data, etc. The ACL specifies what data access is granted to system processes or users and their operations. The operation and a subject are specified in all entries of the ACL. For example, an ACL file has been read, and Alice is given permission to read a specific EHR type. When an EHR type operation is requested by subject X, the ACL is checked by the operating system to determine its legitimacy. However, there are certain challenges with the ACL model, such as how to edit the control access lists, for example, what access is granted to processes and users. An alert will be generated by the system, and the relevant personnel are informed when subject X does not have data access.

Case III: Unauthorized access to patient EHR through the ACL server: In this scenario, subject X attempts to bypass the Access Control List (ACL) server to illicitly obtain specific patient EHR data. The protocol requires robust authorization mechanisms, where X is required to acquire a legitimate authorization key from the authorization server to access any EHR data. This key is a critical component in protecting patient information, as it enables only authorized personnel to retrieve EHR data. The underlying principle emphasizes the importance of strict access controls in preventing unauthorized data breaches, aligning with security frameworks. In this context, the ACL server functions as a gatekeeper, ensuring that access to sensitive EHR data is tightly regulated and complies with HIPAA guidelines.

Case IV: System operators try to abuse patient EHRs: The proposed model uses homomorphic encryption on its database server which encrypts patient EHR data. The advantage of this type of encryption is that patients can modify their information and system operators do not need to have knowledge about this. The operators also do not know in which profiles this modification has occurred.

Case V: Restriction on Patients Accessing Other Patients' EHRs: This case explores the scenario in which a patient attempts to access or allow access to another patient's EHRs. The system's authentication server plays a critical role in this context, rigorously authenticating each patient upon entry into the system. Each patient is assigned a unique private EHR decryption key, which fundamentally prevents them from accessing or decrypting other patients' EHR data. This design principle adheres to the concept of 'Least Privilege' ensuring that individuals have the minimum level of access or permissions necessary to perform their functions. The architecture of this system inherently protects against the possibility that patients compromise the EHR data of other patients, thus reinforcing the security protocols recommended in the healthcare data protection guidelines.

Case VI: Man in the Middle Attack: There is no possibility for a man-in-the-middle attack to occur in this proposed framework. For example, consider cases where patient EHR data have been used or updated. (i) Update in the cloud server: Patient consent, ACL server, and authentication server are in place to ensure access and authentication right to update the patient profile by any EHR user. These servers maintain a session mechanism to protect them from man-in-the-middle attacks. (ii) Display and update at the end of the doctor: If the patient's EHR data is accessed by physicians / specialists, it is assumed that the patients are available for the session. The encrypted EHR data from the server is used by physicians. Then the private key provided by the patients is used to decrypt and use the EHR information. If updates are needed with this information, doctors will use the patient's private key to encrypt the data and then save it on the server. Hence, there is no possibility for a man-in-the-middle attack since the EHR data are encrypted.

2.4.3 What are the main strengths and weaknesses of EHR?

The implementation of EHR represents a transformative change in healthcare care management and patient care around the world. Moving away from conventional paper-based records, EHRs offer a more dynamic, efficient, and accurate way of handling patient health information. Integration of EHR systems is a response to the growing need for coordinated patient-centric care that relies on prompt and accurate exchange of health data. In this section, we will perform a detailed examination of the strengths and weaknesses of EHRs, addressing the multifaceted impact they have on the healthcare sector. The analysis is structured into two distinct parts: The first part explores the various advantages that EHRs offer, such as improved patient care, improved data management, and overall efficiency of healthcare services. The second part critically assesses the

challenges and limitations faced by EHR systems, focusing on privacy concerns, operational complexities, and integration hurdles with existing healthcare frameworks. This structured approach ensures a comprehensive evaluation of electronic health records, highlighting their significant role in modern healthcare, while also acknowledging the complexities involved in their application.

2.4.3.1 Strengths of EHR

The electronic health record (EHR) system, distinct from the Electronic Medical Record (EMR), offers a more comprehensive patient information repository [26]. Its strengths lie in its ability to support detailed analyses of clinical care and subgroups of patients, particularly those requiring palliative care [27, 28, 167–170]. Integration of social determinants into the EHR improves the delivery of high-quality accountable care [27, 32, 171]. Benefits include improved care levels, increased patient safety, simplified processes, and cost reduction [32]. EHRs facilitate better creation, decision-making, and promotion of health policy [172]. The continuous use of the EHR improves communication, quality of care, reduces medical errors and waste, and transforms the healthcare industry into an information-rich sector [173, 174]. The support of EHRs for Decision Support Systems (DSS) and Intelligent Systems (IS) is notable [117]. The HITECH Act significantly increased EHR adoption, improving care quality in the NHS [32, 175, 176]. EHRs have led to early disease diagnosis, reduced medication errors, compliance with care adherence, and reduced costs [177, 178]. Other specific advantages include the management of epidemics, informed decision making, care coordination, patient satisfaction, and evidence-based care advancement [32]. The strengths of SWOT analysis highlight the timely access and storage capacity of information [41].

2.4.3.2 Weaknesses of EHR

Despite the extensive benefits that EHR systems offer, they are not devoid of weaknesses, particularly in the areas of security and their interconnection with privacy concerns. The initial stages of the implementation of the EHR revealed significant challenges related to data input, security, resource allocation, and concerns about cost and ROI [179]. Among these, data security and privacy issues stand out due to their potential to undermine the integrity and confidentiality of patient information.

Security vulnerabilities in EHR systems can lead to unauthorized access, data breaches, and possible misuse of patient information. These issues are critically interwoven with privacy concerns, as both aim to safeguard patient data, albeit from slightly different perspectives. Security measures are mainly focused on protecting data from external

threats and ensuring data integrity, while privacy measures aim to control access to data based on consent and necessity [174].

The complexity of ensuring robust security in EHR systems is highlighted by the lack of system harmony and the occurrence of patient matching problems, which can create loopholes for security breaches. Furthermore, the intricate workflows associated with EHR and the possibility of insecure data storage increase the risk of data compromise, adding to the physician's burden and potentially affecting patient trust in the healthcare system [32].

To mitigate these risks, it is imperative to adopt advanced security measures that can protect against the ever-evolving landscape of cyber threats. This includes enhancing encryption methods, strengthening authentication protocols, and implementing comprehensive access control mechanisms to ensure that only authorized personnel can access sensitive patient information. Addressing these security challenges is crucial not only for protecting patient data but also for maintaining the confidentiality and integrity of EHRs, thus supporting the overall objective of improving healthcare care delivery [151, 180].

Table 2.3 discusses potential solutions to these challenges, emphasizing the need for continuous improvement in security measures as an integral part of the development and management of the EHR system.

2.4.4 What is the difference between EHR Privacy, Confidentiality and Security?

Every health care organization (*for example* hospitals) is responsible for protecting patients' privacy by ensuring that their electronic health records are secure and confidential [181]. The terms privacy, confidentiality, and security tend to be used interchangeably in the existing literature; however, they refer to different individual protections that may overlap but are not exactly the same [54, 182–188]. Ensure that personal information is secure is one of the most important components of securing someone's data. This involves protecting one's privacy, maintaining data privacy, and / or allowing data to remain anonymous [183]. Security breaches often occur not as the result of a sophisticated technical failure, but as the result of a mistake made by someone with authorized access to information [54]. It is also an individual's choice to disclose their problem in front of people [187].

Privacy is sometimes confused with confidentiality and security. The right to confidentiality is based on the fundamental rights to privacy and 'informational self-determination',

TABLE 2.3: EHR Complications and Potential Solutions

| Key Issues | Potential Solutions |
|--|--|
| Redundant credentials | Streamline documentation by identifying and focusing on clinically relevant data points that offer the highest utility in patient care management. Consider the adoption of standardized templates that have been recognized for their efficiency in various healthcare settings. |
| Multiple steps and complicated EHR workflows | Reengineer process flows to align with best practices for clinical workflow optimization. Employ user-centered design principles to create interfaces that reduce cognitive load and administrative burden on clinicians. Delegate appropriate data entry tasks to support staff trained for this purpose. |
| Need for automation with new technologies | leveraging emerging technologies such as natural language processing and machine learning to automate routine tasks. Explore the integration of voice-to-text functionalities and interoperable devices to streamline data capture and entry. |
| Closed EHR software platforms | Advocate for open-platform approaches that allow customization and integration of third-party applications. Embrace models that foster a collaborative ecosystem, prioritizing enhancement of the clinician's user experience and specialty-specific functionalities. |
| Information resting in silos | Promote the development and adoption of industry-wide standards for Application Programming Interfaces (APIs) to facilitate robust data exchange. Emphasize the importance of interoperability as a means of achieving a more holistic patient record and improve continuity of care. |
| Poor user experience | Prioritize the simplification of user interfaces and improve mobile optimization. Recognize the increasing prevalence of mobile device usage and the need for responsive design that adapts to various screen sizes and user contexts. |

which are related to protection of personal data [165]. An operational definition of privacy is the fair and authorized processing and access of personal information [87]. However, confidentiality is a different concept and includes more than data protection rights (Figure 2.4). Firstly, confidentiality works downstream of privacy, and for confidentiality to be legally "triggered", privacy must have already been disclosed. Furthermore, the right to privacy is called a 'negative' right because it claims non-interference with information belonging to the private sphere [101]. Privacy and confidentiality are among the inalienable rights of all human beings that contribute to the preservation of a sense

of reverence and dignity [135]. Privacy refers to an individual's control over how much, when, and under what circumstances they may share details of their physical, behavioral, or intellectual life with others, and their right to restrict other people's access to their personal information [144]. Privacy requirements typically arise in two forms. First, many organizations adopt privacy policies based on their own ethical sense of proper information handling. Second, a variety of laws and regulations impose privacy requirements on institutions and organizations [54]. Data security is the use of logical, technical, administrative, and physical safeguards to ensure the confidentiality, integrity, and availability of data. But confidentiality prevents authorized access to nonpublic information that two or more parties have agreed to restrict [87]. Thus, confidentiality means that providing information to another person will result in a commitment on their part not to reveal it to anyone else [144]. In clinical contexts, hospitalized patients have limitations that may jeopardize their privacy and therefore have serious consequences. [102]. Furthermore, a commitment to confidentiality provides the basis for trust in therapeutic communication.

TABLE 2.4: Clarifying the Components of Information Security

| Component | Definition | Role in Information Security |
|-----------------|---|---|
| Confidentiality | The principle that information should not be made available or disclosed to unauthorized individuals, entities, or processes. | Confidentiality is a key aspect of information security that involves restricting access to information to protect personal privacy and proprietary knowledge. |
| Integrity | The assurance that information is trustworthy and accurate and has not been tampered with or altered by an unauthorized party. | Integrity is crucial for maintaining the trustworthiness of information systems and ensuring that data are not altered in an unauthorized manner. |
| Availability | The guarantee that authorized users have reliable and timely access to information and associated assets when needed. | Availability is essential for ensuring that information systems operate effectively and that data can be accessed when required. |
| Security | A broader term that encompasses the practices and procedures designed to protect information from unauthorized access, use, disclosure, disruption, modification, or destruction. | Security as a whole is built on the principles of confidentiality, integrity, and availability, often referred to as the CIA triad. Each component plays a distinct role in the protection of information assets. |

The privacy issue is one that often applies to the right of patients to protect their information from any other person. It involves the protection of vulnerable data such

as personal data, demographic data, disease symptoms, medical history, test reports, medication record *etc.* from being openly disseminated to others (specialists, radiologists, pharmacists, researchers *etc.*). In general, privacy is the individual's right to keep their data private. Confidentiality is a similar idea, but with a slightly different component. Confidentiality agreements are often applied to situations where someone trusted with personal data must protect these data from being released. Alternately, some studies define confidentiality as issues about the data that get collected, where privacy issues have to do, again, with the core principle of an individual not being recorded or monitored. Security is a different term that is applied to organizational systems. Security may include the idea of customer privacy, but the two are not synonymous. Likewise, security may provide for confidentiality, but that is not its overall goal. The overall goal of most security systems is to protect the healthcare organization, which may or may not house many patient data. Sometimes, the objectives for privacy and security are the same. In other cases, security may not automatically cover privacy concerns. An example is where a healthcare organization may be able to keep its data safe from outside attackers, but where staff (doctors, nurses) may be able to view patient information. Another scenario might involve situations where an organization (*e.g.* hospital), doesn't face any liability by releasing patient data and so chooses to do so. Here, hospital security is not jeopardized, but patient privacy is violated. The studies in [87, 189] describe three areas of overlap between privacy and information security:

- Integrity (information security) and accuracy (privacy): The integrity requirement of information security overlaps with the privacy accuracy requirement in that both are designed to ensure that data are not altered without both authentication and authorization.
- Availability (information security) and access (privacy): Information security availability requirement supports privacy access requirement because if the data is not available, it cannot be accessed.
- Accountability (both): Both information security and privacy doctrines require data owners and custodians to be responsible for protecting data according to the respective protection regimen, which is a form of accountability.

Between privacy and confidentiality, privacy is about personal or private *i.e.* while security and privacy are interdependent, security can be achieved without privacy, but privacy cannot be achieved without security. Security protects confidentiality, integrity, and availability of information, whereas privacy is more granular about privacy rights with respect to personal information. Privacy prevails when it comes to processing personal data, while security means protecting information assets from unauthorized

access. Personal data may refer to any information concerning any individual such as names, addresses, credentials, financial accounts information, social security numbers, *etc.* [189].

To ensure a high degree of privacy and protect user data in the EHR system, various criteria must be met. Privacy advocates and regulators have devised viable strategies that promote privacy protection, similar to the GDPR, which is a collection of rules that gives EU citizens more control over their personal data. Under the GDPR, organizations are obligated to ensure that personal data is collected legally and under strict conditions. Furthermore, those who collect data can be legally liable for any resulting misuse or exploitation in the event of any negligence. The Data Protection Act 2019 works similarly in the USA [190]. Healthcare data have value and are attractive to cyber criminals who wish to inflict data extortion attacks. The overall performance of healthcare systems is affected by this unauthorized access [113, 120, 149].

Many researchers and policy acts have addressed and emphasized the privacy of patient EHRs [7, 8, 10, 102, 104, 105, 122, 123]. The work in [8] illustrates that there is a lack of security in current EHR frameworks, a lack of privacy for patients, and an unreliable method of transmitting their health data, especially in urgent situations. Due to the high need for data privacy, HIPAA and HITECH have implemented many EHR safety protections in the US [156]. The GDPR guidelines have been implemented to ensure data protection and subsequent rights for EU citizens [158]. In Australia, 13 privacy principles have been implemented in relation to the use, disclosure, and sharing of personal information [191, 192]. Although they have local and international privacy policies, EHR systems have faced many data breaches, which resulted in a lack of trust in using existing health record systems.

2.4.5 What different technologies are available to preserve the privacy of EHRs?

A range of information (sharing) security and privacy strategies have been introduced and implemented, but the cloud, NLP, cryptography, and blockchain are seen as the most effective [104, 115, 156, 193, 194]. We divided the various techniques and technologies for preserving EHR privacy into the following four categories: for our review as ABC *i.e.* *Access Control Techniques, Blockchain Techniques, Cloud-based Techniques, Cryptography Techniques.*

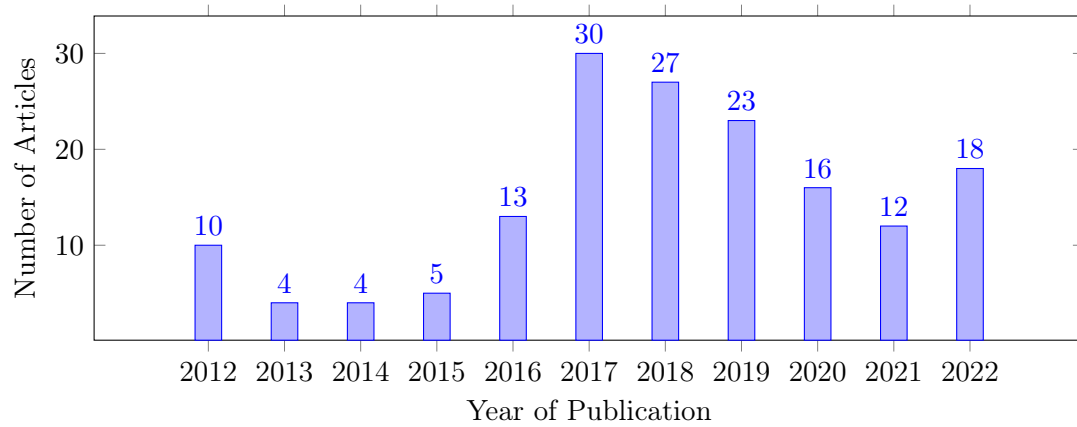


FIGURE 2.3: Distribution of Research Articles from 2013 to 2022

2.4.5.1 Access Control Techniques

TABLE 2.5: Categorization of Technologies, ABC (Access Control, Blockchain, Cloud-based, Cryptography)

| Strengths | Limitations | Included papers | References |
|---|--|--|---|
| Access control | | | |
| <ol style="list-style-type: none"> 1. Limits access rights to unauthorized users 2. Gives patients greater control 3. Reduces administrative overheads 4. Provides granularity of system privilege management | <ol style="list-style-type: none"> 1. reliance on manual input 2. constant need for maintenance 3. too many roles assigned to a person may lead to role explosion causing security holes. | <p>Papers that focus on allowing role-based access to handle various types of users who possess different access privileges. Papers that discuss a hierarchical access structure to grant access to authorized users and limit access rights to other users in the public domain, smart contracts for decentralized data sharing and providing patients with access control over their records and eliminating the need for management services provided by the record-generating parties.</p> | <p>[87, 101, 151, 175, 180, 184, 195–200]</p> |
| Blockchain | | | |
| Continued on next page | | | |

Table 2.5 – continued from previous page

| | | | |
|--|--|---|--|
| <ol style="list-style-type: none"> 1. Immutability 2. Transparency 3. Reliability 4. Interoperability 5. Data provenance | <ol style="list-style-type: none"> 1. Block timestamp dependency 2. Re-entrancy problem 3. Unchecked and failed send 4. No restricted transfer | <p>Papers that focus on blockchain-based strategies for healthcare, the mitigation of problems associated with the privacy and integrity of patient information, the features of blockchain, such as immutability, transparency and reliability, and other factors. Papers also include a blockchain-based framework for storing EHRs, aiming to tackle problems such as response time in data access, interoperability, and better data quality.</p> | <p>[13, 87, 114, 121, 123, 160, 164, 176, 181, 185, 201–205]</p> |
| Cloud-based Techniques | | | |
| <ol style="list-style-type: none"> 1. Large scale and on-demand storage 2. Easy Data recovery 3. Syncing and updating 4. Mobility 5. Quick deployment | <ol style="list-style-type: none"> 1. Downtime due to power failure 2. Provider login 3. Platform dependency 4. High variation in cost due to implementation of additional application | <p>Papers that focus on techniques to ensure the integrity and traceability of medical data over a network, frameworks/models/taxonomies that support the improvement of user security over a network.</p> | <p>[8, 106, 115, 129, 146, 156, 158, 165, 187, 190, 206–208]</p> |
| Cryptography | | | |
| Continued on next page | | | |

Table 2.5 – continued from previous page

| | | | |
|--|--|--|---|
| <ol style="list-style-type: none"> 1. Authorization 2. Consistency 3. Confidentiality 4. No Data Voilation 5. Encryption Is On The Data | <ol style="list-style-type: none"> 1. The size of the key provides a lower bound on the security of the cryptosystem. 2. The hash function can be tempered by two arbitrary inputs that have the same hash value. 3. Information available from physical implementation of cryptosystem can be attacked | <p>Papers that include cryptographic techniques and methods such as: privacy- pre- serving medical record searching scheme for intelligent diagnosis, guarantee a tight data security, securely invoke and share past medical records to make diagnosis. Papers also cover the use of secure searching without leaking any other information on the two parties.</p> | <p>[6, 20, 141, 151, 158, 174, 188, 191, 196, 204, 209, 209, 210]</p> |
|--|--|--|---|

Several EHR services are based on blockchain, cryptography, or cloud (**Table 2.5**). Most of the proposed approaches do not provide an attribute-based access control and encryption mechanism. Various access control models have been proposed, namely *mandatory access control* (MAC), RBAC [11] and ABAC [211, 212]. Modeling access control policies has been a topic of interest. XACML is a policy model based on the XML specification language [185, 213] that uses attributes to impose access control [3, 15].

EHR System with role-based access control(RBAC) [104]: This enables various types of users to have different access privileges. Its hierarchical access structure grants access to authorized users and limits access rights to other users in the public domain. The policy transformation approach enables EHR data to be transferred from a private cloud to a public cloud with the corresponding transformation in the access control policy.

HIPAA three tier themes with respect to administrative, physical, and technical safeguards [156]: These three safeguards encompass a vast array of security techniques that healthcare organizations implement to further secure and protect the health information contained within the EHR. It focuses on compliance with security policies and procedures to prevent physical access to protected health information through unauthorized access to hardware and software by unauthorized users.

ESPAC: This implements granularity authorization for data queries, based on ABE in eHealth [26, 143, 196].

Access control scheme: This is based on elliptic curve cryptography, but there is no support to control access granularity in the proposed authorization process [143, 213].

GAA-FQ (Granular Access Authorization Supporting Flexible Queries): This comprises an access model and an access authorization scheme. Unlike existing blockchain schemes, this access model can authorize different levels of granularity of authorization, whilst maintaining compatibility with the underlying blockchain data structure. Furthermore, the authorization, encryption, and decryption algorithms proposed in the GAA-FQ scheme eliminate the need to use a *Public Key Infrastructure* (PKI) and hence improve the computation performance needed to support more granular and distributed, yet authorized, EMR data queries [143].

In addressing the robustness of security measures for Electronic Health Records (EHR), it is pivotal to acknowledge the role of Multi-Factor Authentication (MFA). MFA introduces an enhanced layer of security by requiring multiple user verification forms before allowing access to sensitive data [105, 147]. This methodology significantly mitigates the risk of unauthorized access by amalgamating various forms of verification: something the user knows (such as a password), something the user possesses (such as a security token or mobile application), or an inherent personal attribute (biometric verification, e.g., fingerprint or facial recognition) [11, 148].

2.4.5.2 Blockchain Techniques

The blockchain, originally proposed in 2008 and used since 2009 [214], is fundamental in the establishment of the Bitcoin network and facilitates non-third-party transactions. Its applications span financial services, reputation management, and the *Internet of Things* (IoT). In healthcare, blockchain is critical for secure data transmission, particularly in the privacy of electronic health records [155, 191], biomedical data [154, 213] and e-health data sharing [11]. Features such as immutability, privacy, transparency, decentralization, and distributed ledgers enhance its appeal [148, 209, 214, 215].

Various scholars suggest the blockchain for increased accuracy, security, and privacy preservation [17, 121, 189, 216]. However, challenges include cultural changes, multiple access nodes, and implementation of a centralized system [138, 146, 217].

Comparative Analysis of Blockchain Solutions:

- *MedRec* and *MediBloc* both utilize blockchain for EHR management but differ in their approach. *MedRec* uses a combination of blockchain for metadata storage and *Distributed Hash Table* (DHT) with *InterPlanetary File System* (IPFS) for actual data, using Ethereum smart contracts for access control [150, 213, 218]. *MediBloc*, however, centralizes the patient as the data flow medium, emphasizing patient-centric data usage and data sovereignty in healthcare care [219, 220].

- *Decentralized Medication Management System (DMMS)* and *Medical Chain* demonstrate the use of blockchain for prescription management and patient access control, respectively. Although DMMS focuses on encrypted prescription sharing [221], Medicalchain uses Hyperledger Fabric for patient-controlled data access [197].
- *Blockchain-Based Data Sharing Mechanism* [101] and *Blockchain-based searchable encryption* [155] highlight the role of blockchain in secure data sharing and accurate search results in research networks, focusing on cryptographic algorithms and data integrity.
- *Estonia's healthcare system* and *Healthcare Data Gateways (HDG)* integrate blockchain for data integrity and security. Estonia's system uses *Keyless Signature Infrastructure (KSI)* blockchain [203, 222], while HDG focuses on access granularity and attribute-specific data queries [143, 205].

The consensus algorithms used in these solutions vary. For instance, Ethereum-based systems like MedRec typically employ Proof of Work (PoW) or transition to Proof of Stake (PoS), while Hyperledger Fabric used in Medicalchain and OmniPHR opts for more customizable consensus mechanisms.

Incorporating these comparative aspects provides a subtle understanding of blockchain's versatility and adaptability in healthcare, particularly in enhancing EHR privacy and security.

Taxonomy of Blockchain systems:

- **Public blockchain:** Open to anyone to join, such as Bitcoin [123] and Ethereum [149, 195].
- **Private blockchain:** Requires invitation or authorization, e.g., MultiChain [121] and GemOS [151].
- **Consortium blockchain:** Semi-private, used by authorized organizations, exemplified by Hyperledger Fabric [216] and Ethereum for consortium blockchains.

Blockchain, while promising, has limitations such as slow processing, scalability issues, privacy challenges, and high energy consumption [147]. Understanding these subtleties is crucial for the advancement of blockchain applications in healthcare, particularly for the privacy of EHR.

Blockchain based strategy A Survey of Blockchain-Based Strategies for Healthcare: A blockchain-based strategy can mitigate problems arising from issues threatening

the privacy and integrity of patient information, due to blockchain's immutability, transparency and reliability [215].

MedRec [150]: This is a blockchain-based framework for storing EHRs that aims to address problems such as response time in data access, interoperability, and data quality. MedRec [218] is a blockchain-based decentralized record management system to handle EHRs. Meta-data are stored on blockchain and the real data is stored on *Distributed Hash Table* (DHT) by using *Inter Planetary File System* (IPFS). A smart contract is used for access control and there is a transaction fee [180, 202].

MedShare: A MedShare-based solution involves a system consisting of four layers: (i) User layer: the data will be accessed through a graphical interface; (ii) Data query layer: a group of structures that process and respond to query requests in the system; (iii) Database infrastructure layer: a layer composed of the system databases, to which only a few specialist institutions have access; (iv) Data structuring and provenance layer: responsible for processing within the system; in other words, it is the layer that contains the adopted blockchain network structure, consensus protocol, node authenticator, and smart contracts. It offers features such as data provenance, auditing, and greater security for systems [223].

Medicalchain: This was built with the help of a permissioned blockchain from Hyperledger Fabric. The application enables patients to have access controls for all their information, as well as being able to handle their healthcare data in a personalized way [197].

MediBchain: : A novel blockchain-based EHR automation system for healthcare. It is a patient-centric healthcare data management system that uses blockchain as storage to maintain privacy. A decentralized feature of blockchain technology is that it eliminates vulnerabilities to protect data and maintain privacy and security [224, 225].

Decentralized Medication Management System (DMMS): A novel blockchain-based EHR automation system for healthcare. A physician examines the patient and writes a prescription. The prescription is encrypted with the patient's public key and no one can access the patient's record without their private key. The patient can view his record and, at the same time, the doctor can also view the patient's record with the approval of the patient [221].

Healthcare Data Gateway app: This is a blockchain-based security & privacy system for biomedical and healthcare. Information exchange systems enable patients to own, control, or share data securely without infringing privacy, offering a new way to improve healthcare systems while maintaining patient data confidentiality [205].

Blockchain-Based Data Sharing Mechanism [101]: This provides a secure distributed research data sharing network and a way to specify/control the parameters of sharing, providing full accountability of access to such data. **Blockchain based searchable encryption** [155]: This guarantees that the data user can receive accurate search results without additional verification. It enables cryptographic algorithms to be built to ensure data integrity, standardized auditing, and some formalized contracts for data access.

Decentralized and Hierarchical Data Sharing using Smart Contracts [8]: This scheme empowers patients by giving them control over their records, allowing them to selectively share data with users that satisfy their privacy preferences. Give patients access to control over their records and eliminate the need for management services provided by record-generating parties.

Estonia health care system and Personal Care Record Platform (MyPCR): It is related to health data with its requirements, challenges, and existing techniques for data security and privacy. It use *Keyless Signature Infrastructure* (KSI) blockchain to ensure the integrity and security of the data in its system [203, 222].

Healthcare Data Gateways (HDG): Its access granularity is based on blocks. It cannot support data queries to specific data attributes in blocks or restrict access authorization to these attributes [143, 205].

Ancile, Privacy-preserving framework for access control and interoperability of EHR using blockchain technology [11]: A blockchain-based framework for secure, interoperable and efficient access to medical records by patients, providers, and third parties while preserving the privacy of patient sensitive information. This framework, Ancile, utilizes smart contracts in an Ethereum-based blockchain for increased access control and data obfuscation and employs advanced cryptographic techniques for greater security.

Decentralized and Hierarchical Data Sharing Using Smart Contracts [8]: This is a decentralized blockchain technology to mitigate security issues, privacy concerns, and the inefficiencies of various centralized platforms such as financial systems. It is a secure, private, and efficient electronic record sharing scheme that utilizes smart contracts deployed over a blockchain.

MediBloc [219]: This is an open source healthcare data platform built on blockchain that can secure and integrate diffused data from numerous organizations. It can track a person's daily movements using smartphones, fit bands, smart watches, etc. [204] but it has performance, scalability, and energy consumption issues. Medibloc also uses meta-data, however, the operations are different. In MediBloc, patients are the medium of

data flow and exchange and utilize their health data as needed [220]. Integrates multiple hospital records into one secure decentralized ledger to establish a medical record data base on blockchain. MediBloc uses a public blockchain that allows anyone to access transactions. Minimize the risks of personal healthcare information leakage and maximize the credibility of medical records. It also provides reliable, personalized, and patient-centric health information. MediBloc ensures the privacy of health information and improves data sovereignty in the medical ecosystem [57].

OmniPHR [226] uses the concept of blockchain (linking blocks) to store data that is broken into small pieces as blocks. The system improved interoperability, storage, and scalability. The data stored on the blockchain are encrypted with a key that is generated and stored by the body sensor node. This system can only be used for body sensor networks (wearable devices); however, PHRs include health data from various resources.

MedVault [147] also stores health care data on blockchain and is a privacy-preserving system. It is an attribute-based authentication system that enables EHR sharing in a patient-centric manner. But their study results showed that MedVault performed well supporting all types of EHR subjects, but not with patient and physician subjects. This is mainly due to the fact that MedVault considers EHR accessibility on only all data sets, data elements, and transactions. In addition, it ignores the non-exposure of patient data. [160, 212].

Blockchain-Based Deep Learning as-a-Service (BinDaaS): is an architectural framework for secure transmission of EHRs [199, 201]. Integrates blockchain and deep learning techniques to share EHR records between multiple healthcare users.

TABLE 2.6: Comparative Analysis of Blockchain-based Solutions for EHR

| Solution | Blockchain Technology | Consensus Algorithm |
|---|-----------------------|--|
| MedRec [150, 180, 202, 218] | Ethereum | Proof of Stake (PoS) |
| MedShare [223] | Custom Blockchain | Not Specified |
| Medicalchain [197] | Hyperledger Fabric | Practical Byzantine Fault Tolerance (PBFT) |
| MediBehain [224, 225] | Custom Blockchain | Not Specified |
| Decentralized Medication Management System (DMMS) [221] | Custom Blockchain | Not Specified |
| Healthcare Data Gateway app [205] | Ethereum | Proof of Work (PoW) |
| Blockchain-Based Data Sharing Mechanism [101] | Custom Blockchain | Not Specified |
| Decentralized and Hierarchical Data Sharing using Smart Contracts [8] | Ethereum | Smart Contracts |
| Estonia healthcare system and MyPCR [203, 222] | KSI Blockchain | Not Specified |
| Healthcare Data Gateways (HDG) [143, 205] | Custom Blockchain | Not Specified |
| Ancile [11] | Ethereum-based | Smart Contracts |
| MediBloc [57, 204, 219, 220] | Public Blockchain | Not Specified |
| OmniPHR [226] | Custom Blockchain | Not Specified |
| MedVault [147, 160, 212] | Private Blockchain | Practical Byzantine Fault Tolerance (PBFT) |
| Blockchain-Based Deep Learning as-a-Service (BinDaaS) [199, 201] | Custom Blockchain | Not Specified |

Table 6 presents a concise comparative analysis of key blockchain-based solutions in the EHR domain. Highlights various approaches adopted by these solutions, categorizing them according to their distinct blockchain technologies and consensus algorithms. This table is an essential tool for readers to understand the diverse applications of blockchain in EHR, particularly to enhance privacy and accessibility. By detailing the technical

foundation of each solution, the table helps to understand the critical factors that contribute to the security, scalability, and effectiveness of these solutions in managing EHR privacy and accessibility concerns.

2.4.5.3 Cloud-based Techniques

Cloud-based platforms are useful for delivering electronic health services with ubiquitous network access, scalability, and cost savings. Transferring electronic health records to the cloud poses major threats to privacy, data integrity, and confidentiality, and additional techniques are required to maintain data secrecy. Cloud-based utility services (such as storage) offer additional benefits to EHR systems: for example, they are more cost effective, can be easier to manage (for example, access and retrieval), and support collaboration, with mobile technologies and devices to gather data [217].

Cloud-based EHR system Using Attribute-Based Cryptosystem and Blockchain [115, 227]: Wang and Song [10] proposed a cloud-based EHR system that uses ABE and IBE to encrypt data, ensuring fine-grained access control for encrypted data using an *identity-based signature* (IBS) to implement digital signatures.

Attribute Based Encryption for Secure Access to Cloud-Based EHR Systems [15]: Through this system, every patient's visit is recorded as a separate node on the knowledge graph, ensuring strict data security, making it easier to query and speed up data access procedures.

2.4.5.4 Cryptography Techniques

To prevent unauthorized users from accessing EHRs, a direct way is to encrypt EHRs before uploading them to cloud servers [159]. To protect data privacy and mitigate threats, various encryption models have been proposed. ABE is one such interesting approach where the ciphertext, the secret key, and the private key of the user are associated with the user attributes [16, 101, 150]. Bethencourt et al. [228] developed a system called *Citro-Policy Attribute-Based Encryption* (CPABE) to implement ABE using user attributes to encrypt the document [101].

Cryptographic Role-Based Access Control Model: This ensures secure access to EHR resources by enforcing cryptographic access control with context and location awareness. A role-based cryptographic access control model for EHR systems uses location- and biometrics-based user authentication and a steganography-based technique to embed EHR data in host electrocardiographic signals [217].

My Health My Data (MHMD): This technique is for data security and privacy sharing of medical information and for empowering its primary owner, the patient [229–231].

Ancile: Privacy-preserving framework: : Using six separate contracts, Ancile improves the efficiency of the patient experience and reduces privacy threats. The patient is the only node expressly given the location of their information. Ancile maintains the cryptographic hashes of stored records and query links, which confirm the integrity of EHR databases.

Keyword searchable encryption and proxy re-encryption technology [115, 147]: Protect data security with a searchability guarantee that only authorized entities can access EHRs. [115] combines keyword searchable encryption and proxy reencryption technology to ensure privacy preservation and secure data sharing for EHRs based on consortium blockchain technology and cloud storage. The secure cryptographic technique (proxy re-encryption) is applied to support efficient access control on secret data [147]. Re-encryption of cyphertext by the cloud ensures relatively good security to the data using the technique proposed by [115]. But the keyword searchable encryption method is not clearly described from the user’s point of view. This technology can be adapted to make data in the cloud more secure and is capable of identifying and giving access to the right user [115, 147].

Privacy-preserving medical record searching scheme (PMRSS): This is a scheme for intelligent diagnosis in IoT healthcare that securely invokes and shares past medical records to assist in a diagnosis. The input used for the search must be protected as well as the result. It securely searches the diagnosis report by only two rounds of interactions without leaking any other information from the two parties.

Personally Controlled Electronic Health Record (PCEHR) System: It uses *Fully homomorphic encryption* (FHE) to encrypt patient data. The decryption key is held by the patient; therefore, no other person can access the data without the patient’s permission. [12] uses a verification technique, such as cryptography, to ensure that only an authorized person can access the corresponding records [15, 232].

2.4.6 Enhancing EHR Security with Multi-Factor Authentication

Multifactor authentication (MFA) plays a crucial role in fortifying the security measures surrounding Electronic Health Records (EHRs). By requiring users to provide multiple forms of verification before granting access to sensitive information, MFA significantly reduces the likelihood of unauthorized data breaches [105, 147]. This approach typically

combines two or more independent credentials: something the user knows (password or PIN), something the user has (security token or authentication app), and something the user is (biometric verification like fingerprints or facial recognition) [11, 148].

The integration of MFA into EHR systems, particularly those based on cloud and blockchain technologies, provides an essential layer of security. It ensures strict compliance with stringent privacy regulations, such as the General Data Protection Regulation (GDPR), by meticulously controlling access to medical records [158]. Implementing MFA in conjunction with existing cryptographic and access control measures further strengthens the security framework of EHR systems, protecting against unauthorized access and enhancing trust in digital healthcare services.

2.5 SURVEY FINDINGS

After examining the selected literature and answering the survey questions, we identified some major points.

2.5.1 Privacy, Confidentiality and Security: the Difference

The terms privacy, confidentiality and security are used interchangeably as they refer to related concepts. But there is a need to address the inconsistent usage of such terminologies as they actually have varied definitions. Data security governs access to data throughout the data life cycle. In contrast, data privacy sets this access based on privacy policies and laws that determine, for instance, who can view personal data, financial, medical or confidential information. The three main concepts of security are authentication, authorization, and access control [125, 160, 233, 234]. Therefore, we can say that confidentiality protects secrets, security is broader than confidentiality, and privacy determines authorization. Privacy is closely related to security and confidentiality, but approaches data from a different perspective. Confidentiality controls and protects against the unauthorized use of information already in the hands of an institution, whereas privacy protects the rights of an individual to control the information that the institution collects, maintains and shares with others. One way to understand the relationship between privacy and confidentiality is that privacy requirements dictate the types of authorization granted to information, and confidentiality controls ensure that people and systems meet those privacy obligations (**Table 2.7**). Therefore, when it comes to EHR management, it is important first to understand the difference between its security, confidentiality, and privacy. It is at one's (or the organization's) risk to substitute one for the other. There is a rich set of tools with which to protect EHRs, but

TABLE 2.7: Difference Between Privacy, Confidentiality and Security

| Basis for Comparison | Privacy | Confidentiality | Security |
|-----------------------------|--|---|--|
| Meaning | The state of being free from intrusion or interference | A situation where someone is not expected to divulge information to any other person. | The state of being free from danger or threat. |
| What is it | Right to be let alone. | Agreement between the persons acting as fiduciaries to maintain the secrecy of sensitive information and documents. | the process and practice of safeguarding data throughout its entire life-cycle. |
| Concept | Limits the access of the public. | Prevents information and documents from unauthorized access. | Protect data from unauthorized access, corruption, theft, damage, or loss by implementing specific controls, standard policies, and procedures |
| Applies to | Individual | Information | Organization / System |
| Obligatory | No, it is the personal choice of an individual | Yes, when the information is professional and legal. | Yes, data are a valuable asset. |
| Disallowed | Everyone is disallowed from being involved in the personal affairs of an individual. | Only unauthorized persons are disallowed from using the information | Every harmful activity |

it is important for all data protection practitioners, IP attorneys, information security specialists, and privacy professionals to be aware of the health records in question and ensure that the proper protection paradigm is applied.

2.5.2 EHR Privacy Concerns

This review focused solely on the privacy aspect of the EHR, that is, how patient records are kept private under various circumstances and what different techniques the researchers use to ensure the privacy of the patient's EHR. Technologies can host several risks; hence, the privacy of information in these systems is of utmost importance. Regardless of the increased effectiveness and growing interest in the use of EHRs, little attention is paid to privacy issues that might arise. Mobility and the use of multiple mobile devices in collaborative healthcare increases the need for robust preservation of privacy. Therefore, large-scale EHR systems require secure access to sensitive data, data storage, and management [104]. One of the major security concerns is the issue of the increasing size of healthcare data, but none of the articles reviewed highlighted this matter. As described in [143], HIPAA does not dictate the ways in which to create and implement the systems

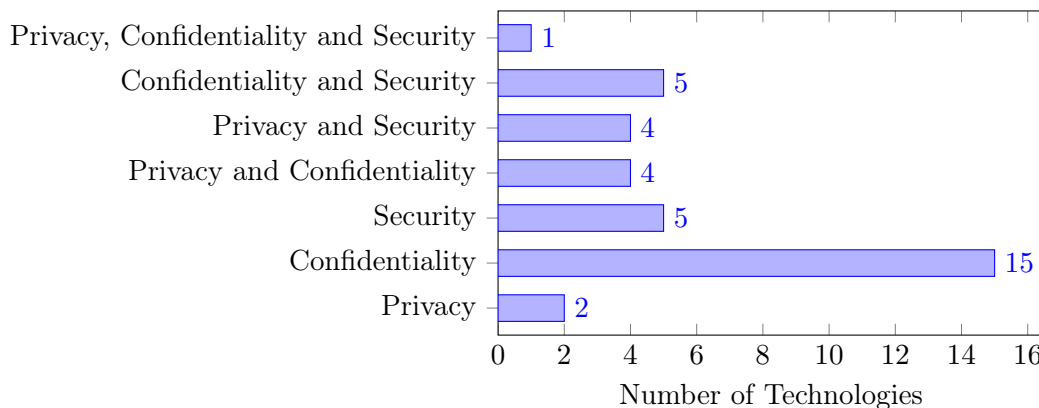


FIGURE 2.4: Number of (Reviewed) Technologies Covering the Aspect(s) i.e., Privacy, Confidentiality, Security

currently being used. This leads to many variations in the centralized systems used today and has prevented interoperability between medical institutions. A major downside of these models is that evidence is not inspected in a structured manner, and regulation procedures and guidelines also differ significantly throughout territories depending on local experience and national implementation of privacy policies [57, 233, 235]. Table 2.8 lists the different technologies related to privacy, confidentiality and security. Privacy preservation must be reviewed in light of changing privacy rules and legislation on sensitive personal data. Users must own and control their data without compromising security or limiting the ability of companies and authorities to provide personalized services. Researchers [143] are in favor of blockchain, smart contracts, and their implementation by Ethereum to enforce verified negotiations of contracts between two participating parties over the blockchain. Like any other transaction processed over the blockchain, they are based on cryptographic primitives that ensure their integrity. Some researchers believe that both the cloud and blockchain can be used in combination to provide cost-effective security solutions, but organizations have realized that a one-size-fits-all approach may not work for cloud adoption in the case of public and private clouds.

Access control: Several studies proposed solutions for privacy-preserving data sharing based on ABE or CP-ABE in the cloud to encrypt data and provide the hierarchical access structure for fine-grained data sharing. However, they did not provide policy dynamics. One of the challenges of data sharing is key management. Yu et al. [13] identified the data security and access control issues associated with the sharing of EHRs within the public domain due to the high computation overhead in key distribution and data management, which occurs when applying fine-grained access control. They used *Key-Policy ABE* (KP-ABE), *Proxy Re-Encryption* (PRE), and lazy re-encryption to define and enforce access control policies, but implementing secure and dynamic access rights is challenging [226].

Blockchain: The work in [103] proposed a blockchain-based framework to manage, maintain and share electronic medical records from cancer patients. They adopted permissioned blockchain technology to access, manage, and store encrypted patient data. Such proposed frameworks can be used to practically implement blockchain technology to access and manage the privacy and security of patient data and history in clinical practices. The Ancile framework discussed in [11] used smart contracts and permissioned blockchains, but it is still in the early stages of development in the Ethereum community. However, one cannot rely solely on Ancile as a remedy to the larger EHR security problem, but it can be adapted and incorporated into another technique to achieve optimum results. The work in [147, 233, 236] described various algorithms to efficiently share EHRs in blockchain-based electronic healthcare record systems for healthcare 4.0 applications with less communication time. The algorithms proposed in their article cover the maximum number of collaborating parties that could be involved in EHRs. The security of the proposed system is evaluated by its performance through simulations and scenarios which is missing from other proposed approaches. However, they only considered admin, patients, clinicians and laboratory personnel as participants in the EHR system, while other multiple participants *such as*, health organizations can be involved.

During the implementation of an EHR, the sharing of medical data often faces critical limitations, such as loss of control over the data, provenance of the data, auditing, and safe data trailing in medical data. To address these limitations, MeDShare provides a secure and safe blockchain system for the exchange of medical data among untrusted parties. MeDShare can be used to share medical data and maintain EHRs among cloud service providers, hospitals, and healthcare research entities with greater data provenance, personalized audit control, and minimal possible threats to data security and privacy [205].

Cloud-based: In cloud-based EHR, the dissemination of patient data is greatly beneficial, but must be done in such a way that patient privacy is preserved. The model proposed in [143, 236] also follows a patient-centric approach to EHR management where the responsibility of authorizing data access is handled at the patient's end. However, this creates significant overhead for the patient, who must authorize every access to his health record. This is not practical given the multiple personnel involved in providing care and that, at times, the patient may not be in a state to provide this authorization. Hence, there is a need to develop a proper authorization delegation mechanism for secure, secure, and easy cloud-based EHR management. Despite existing solutions, privacy issues are the major obstacles that limit the widespread adoption of public clouds across the world. The main reason for this concern is that information needs to be published to a broad and possibly anonymous set of receivers, and it can be dangerous to

outsource sensitive data to the cloud, so there is an increasing need to investigate data anonymization techniques applied to this domain.

Cryptography: The ciphertext data based on traditional encryption mechanisms make the sharing of EHR difficult to a large extent. In particular, it is very challenging for resource-limited IoT devices to perform burdensome computation tasks for fine-grained data sharing in mobile cloud computing. To fill this gap, ABE can be adopted to perform fine-grained access control in EHRs [14, 180, 232]. Along with our review of the literature, [135] also points out firewalls and cryptography as the security techniques mentioned most frequently in the selected sample. Various techniques have been introduced to protect patients' privacy by applying various hybrid and cryptographic access control techniques [6, 12, 15, 57, 121, 123, 130, 131, 135, 140, 159, 160, 163, 164, 166, 172, 183, 201, 214, 222, 230, 233, 236, 237]. However, most of these approaches have certain shortcomings that make them less effective with respect to EHR privacy.

Standards' Compliance: The recent cloud-based blockchain approaches suggested by multiple researches [6, 14, 115, 121, 129, 140, 146, 180, 184, 185, 203–205, 233] focus on the implementation of blockchain along with some standards. But the paper contains conflicting descriptions about security standards. First, they favor GDPR subject to data rights and criticize HIPAA about medical records regulations and protecting only PHI but later describe how some GDPR articles directly conflict with blockchain. Therefore, it gives a confusing impression of whether to use GDPR with blockchain. Importantly, distributed methods for data integrity validation are not sufficient to solve all cybersecurity hazards. Despite having much potential to achieve data security for EHRs, existing approaches require further strengthening by complying with the standards *e.g.* HIPAA measures [41, 131, 233] to achieve data privacy, security, and integrity while in a centralized setup.

For medical practices dealing with sensitive patient data, which are required to comply with the US HIPAA rules, a private cloud may be appropriate. The research in [140] also mentions that numerous security standards have been developed, such as HIPAA, COBIT, and DISHA, which have been applied to protect patient health information and can address privacy issues.

2.5.3 Discussion

When discussing EHRs, firstly, it is crucial to first have a technical understanding of the actual definition and characteristics of PCS. Second, every technology has its merits and demerits, so depending on what is needed, the respective technology can be adopted. No limitations have been found in the literature on the use of a single technique to preserve

the privacy of EHRs, but the advantages of combining two or more techniques can be achieved to achieve the desired requirements. On the basis of this review, it can be clearly seen that no technique / solution can be considered optimum for EHR privacy. All techniques utilize different technologies *that is*, cloud computing, Ethereum-based blockchain, cryptography and encryption techniques, and / or access control techniques to ensure data privacy.

2.5.4 Limitations

Despite our best effort to survey as many relevant articles as possible, we present the limitations of this survey. A fundamental constraint identified during our survey process was a general lack of literature that discussed privacy preservation without confusing it with confidentiality and security. As a result, there is a lack of primary articles that compare and contrast the privacy of the EHR with confidentiality and security. Therefore, it was difficult to find techniques and technologies that cover EHR privacy. We also found it challenging to validate some studies simply on the basis of their manuscripts. To our knowledge, none of the existing studies tested their proposed method using either real samples or raw data from EHRs, putting the external validity of these studies in question.

TABLE 2.8: Summary Of Reviewed Technologies And The Aspects They Covered

| Technique | Technology | | | | Management Aspect | | |
|---|----------------|------------|-------------|---------------|-------------------|-----------------|----------|
| | Access Control | Blockchain | Cloud-Based | Cryptogra-phy | Privacy | Confidentiality | Security |
| 1. EHR system with role-based access control (RBAC) | ✓ | | ✓ | | | ✓ | |
| 2. HIPAA implements three safeguards: administrative, physical, and technical | ✓ | | | | | ✓ | ✓ |
| 3. ESPAC | ✓ | ✓ | | ✓ | | ✓ | ✓ |
| 4. access control scheme | ✓ | ✓ | | ✓ | | ✓ | |
| 5. GAA-FQ (Granular Access Authorization Supporting Flexible Queries) | ✓ | | | ✓ | | ✓ | |

Continued on next page

Table 2.8 continued

| | | | | | | | | |
|-----|--|---|---|--|---|---|---|---|
| 6. | Ciphertext-policy attribute-based encryption (CP-ABE) | ✓ | | | ✓ | | ✓ | |
| 7. | Blockchain based strategy | | ✓ | | ✓ | ✓ | ✓ | |
| 8. | MedRec | ✓ | ✓ | | ✓ | | ✓ | |
| 9. | MedShare | | ✓ | | ✓ | | | ✓ |
| 10. | Medicalchain | ✓ | ✓ | | | ✓ | ✓ | |
| 11. | MediBchain | | ✓ | | | ✓ | | ✓ |
| 12. | Decentralized Medication Management System (DMMS) | | ✓ | | | | ✓ | |
| 13. | Healthcare Data Gateway app | | ✓ | | | ✓ | ✓ | ✓ |
| 14. | Blockchain-Based Data Sharing Mechanism | | ✓ | | | | ✓ | ✓ |
| 15. | Blockchain based searchable encryption | ✓ | ✓ | | ✓ | | ✓ | |
| 16. | Decentralized and Hierarchical Data Sharing | ✓ | ✓ | | | | ✓ | |
| 17. | Estonia health care system and Personal Care Record Platform MyPCR | ✓ | ✓ | | | ✓ | | ✓ |
| 18. | Healthcare Data Gateways (HDG) | ✓ | ✓ | | | | ✓ | |
| 19. | Ancile: Privacy-preserving framework | ✓ | ✓ | | ✓ | | ✓ | ✓ |
| 20. | MediBloc | | ✓ | | | ✓ | ✓ | |
| 21. | OmniPHR | | ✓ | | ✓ | | | ✓ |
| 22. | MedVault | ✓ | ✓ | | | ✓ | | |
| 23. | Blockchain-Based Deep Learning as-a-Service (BinDaaS): | | ✓ | | | | | ✓ |

Continued on next page

Table 2.8 continued

| | | | | | | | | |
|-----|--|---|---|---|---|---|---|---|
| 24. | Cloud-based EHR system Using Attribute-Based Cryptosystem and Blockchain | | ✓ | ✓ | ✓ | | ✓ | |
| 25. | Cloud-based EHR system: | ✓ | | ✓ | ✓ | | | ✓ |
| 26. | Attribute Based Encryption for Secure Access to Cloud Based EHR Systems | | | ✓ | | | | ✓ |
| 27. | Cryptographic Role-Based Access Control Model | ✓ | | ✓ | ✓ | | ✓ | |
| 28. | CureMD | | | ✓ | | | | ✓ |
| 29. | Practice Fusion | | | ✓ | | | ✓ | |
| 30. | Athenahealth | | | ✓ | | | ✓ | |
| 31. | MyHealthMyData (MHMD) | | ✓ | | ✓ | ✓ | | ✓ |
| 32. | Ancile: Privacy-preserving framework: | | | | | ✓ | | |
| 33. | Keyword searchable encryption and proxy re-encryption technology | | ✓ | ✓ | ✓ | | ✓ | ✓ |
| 34. | privacy-preserving medical record searching scheme (PMRSS) | | | | ✓ | | ✓ | |
| 35. | Personal Controlled Electronic Health Record (PCEHR) System | | | | ✓ | | ✓ | |

2.6 Conclusion and Future Research Directions

This chapter provides a comprehensive overview of EHRs and their privacy. Although EHRs are widely recognized and recognized for their importance, the survey revealed a lack of systematized knowledge on this topic with respect to their management. Existing surveys dealt with EHRs from more technical and technological perspectives using the three terms interchangeably *that is*., PCS and creates confusion. Instead, the survey presented in this chapter, through the analysis of the reviewed papers, clearly differentiates these three terminologies and gives answers to the research questions, namely (1) the currently available data sharing methods, (2) the role of privacy when sharing EHRs with different stakeholders, (3) the strengths and weaknesses of EHRs, (4) the difference between PCS and (5) the different technologies available to preserve the privacy of EHRs. We believe that future research to protect EHRs should ensure all aspects of privacy, confidentiality, and security. These analyzes supported the identification of future challenges that should drive research in the next few years to obtain a more systematic view of EHR management with the need to clarify concepts specifying the management of EHR , *that is*., PCS. This survey paves the way for a deeper understanding of EHR management beyond technical aspects, contributing to its management by first focusing on requirements. An important aspect that this survey highlighted is that there is no clear definition on the terms privacy, confidentiality, and sec for EHRs. This deserves special attention in establishing a common basis in the study of the differences and similarities of these three terms from the point of view of users and developers. It should be noted that today the utilization of technologies for data management with respect to privacy and security is different from the past, not only because of the growing number and variety of techniques, but also because various techniques can support each other and can be combined to maintain EHR data. Therefore, to preserve the privacy of EHRs, researchers and practitioners must consider the wise and appropriate use of terminologies (*i.e.* privacy, security and confidentiality) and technologies when developing and managing EHR systems, organizational processes, and everything that involves personal health data.

Chapter 3

A SECURITY AND PRIVACY COMPLAINT DATA SHARING SOLUTION FOR HEALTHCARE DATA ECOSYSTEMS: CEMPS (CENTRALIZED EHR MODEL TO PRESERVE PRIVACY AND SECURITY)

3.1 Introduction

The healthcare industry is one of the largest emerging industries globally. This industry involved various types of business, such as offering medical services, manufacturing various medical drugs or equipment, offering health insurance, or facilitating healthcare care provisions to patients. Healthcare industries have seen improvements after integration with technology. Currently, both are highly interconnected. This relationship has been observed over the past decade. The primary reason for this is the easy integration of various devices in all health care centers. For example, telehealth (or e-medicine) was developed before a few decades but is now extremely essential. Helps to distribute health-related information and services to patients who need it through technology. Patient care involves the method of taking care of patients together with their families in a way that provides valuable and meaningful care to an individual patient. This patient care includes involving, informing and listening to patients. According to Baker [238], patient care is defined as providing care that respects and responds to individual patient preferences, needs, and values and ensures that patient values guide all clinical decisions.

Therefore, the latest technologies are needed in the health care sectors to deal with various modern health problems. The primary issue for healthcare sector security developers from an operational and technical perspective is the privacy of patients. The analysis of Jacq [239] identified that when the infrastructures of heterogeneous and large sectors of health information technology do not have the proper implementation of legal requirements, there is no scope for these types of technologies. Additionally, mandatory access

control forms are needed to accommodate the increasing complexity of electronic patient records. The various health records of every patient should be accessed by different types of medical care personnel, such as specialists, general practitioners, nurses and administrative personnel. Pseudonymized or anonymized records are needed for epidemiology or other research purposes. Furthermore, due to the potential for data reidentification, various difficulties and issues are caused by genetic data management. As these are made to maintain simple access rules, consent problems indicated that there is a need for certain fine-grained rules. If patient records are encrypted, a suitable employee should use the decryption algorithm, and audit and recordkeeping facilities must be involved. To prevent sensitive data from being passed to third-party users, an intelligent audit system should be used. However, traditional methodologies are needed to work with modern methodologies to provide the latest architecture models. Data privacy is the main concern with these latest technologies. Data privacy in distributed medical research and health care systems depends on policies directed by legislation and jurisdictions [18, 240, 241]. There is a need to enforce such policies at the program level, but there is no adequate guarantee of privacy protection [19]. However, it is extremely necessary to inspect privacy agreements to improve the social acceptance of health care systems. The studies by Rahmouni et al., [242] and Rahmouni et al., [243] showed the way of semantic web technologies to classify resources. These resources were defined through metadata captured from the data protection and privacy ontology that were implemented and designed. Rahmouni et al., [244] described past model extensions with relevant metadata and also included data sharing situations for suitable healthcare or medical applications. This study also showed how the above study has better authorization and security policy enforcement and specification for cloud computing applications, as explained in the following work of Belaazi et al., [245] and Belaazi et al. [246]. A sensitive data detection model was developed by Essefi et al., [247] for the business process involved in the hospital. The governance of patient data management is simple while applying the rules of HIPAA and GDPR, as discussed by Rahmouni et al. [248]. Furthermore, Munir et al., [111] presented ontology-related query details with the help of assertion and semantic OWL-DL capabilities. This study dealt with the development of highly trustworthy and self-disciplined healthcare sectors through the integration of privacy audit dimensions into patient data management services.

Various works have been developed to control privacy and security problems with EHR. However, many of these works are based on cryptographic approaches and/or the control access approach. The safest method is the cryptography approach to preserve the security and privacy of cloud applications. For the safest data transfer in cloud applications, this cryptography approach is sufficient to practice the public structure key [249–251]. The cryptography approach is used to encrypt sensitive private data, such as clinical

records, before transferring or saving them in cloud storage. Some models were recommended by Li et al. [252], Benaloh et al. [253], Huang et al. [254] and Jin et al. [255] that let patients encrypt medical data before transferring and saving on cloud storage to control potential privacy risks. Public Key Infrastructure (PKI) authentications and digital signatures have been used by Van der Haak et al. [112] to meet the legal requirements of electronic medical record (EMR) exchange. Ateniese et al. [256] proposed a pseudonymization technique that is used to preserve the anonymity of patients. Layouni et al. [257] proposed a communication approach for health monitors to observe the exchange of health information. However, saving sensitive medical data is safer with this cryptography method; accessing those data is difficult. Therefore, the key challenge for EHR applications is the right access to data with this cryptography method [258–261]. To achieve these objectives, incorporating health mechanisms along with providing better privacy in Australia helps people in this country gain advantages over better health and medical systems and strategies with little concern for the privacy of their data [262]. Hence, it is much needed for the adoption of mechanisms that ensure the robustness of privacy implementation proofs and processes to the systems which manage patients' data and offers a better guarantee to such implementation. This can be achieved by auditing and monitoring previous records of healthcare data exchange. However, the privacy of patient data is still at risk despite all automation, especially in the healthcare industry. This can make use of exception-based data access that allows patients to bypass the controls of a system due to unforeseen events or emergency cases [263–265]. Furthermore, these systems can be attacked by intruders or hackers while exposing themselves to cyber security risks and the Internet [266]. Sometimes, failures of a system might be the reason for the risks. Therefore, the implementation of governance and audit mechanisms is needed for the development of healthcare-related applications along with technological aspects. On this basis, the user guidelines and the initial constraints to test conformance during the control access run-time should be matched. This included reviewing the previous data-sharing events of each user and the permissions granted. However, such reviews are hard and expensive when an automated solution is not available. The machine learning approach is also one of its kind and provides various benefits for healthcare data [267–272]. Researchers have used such techniques to protect data and comply with regulations [273–278].

This chapter discusses details of the critical aspects of securing and preserving privacy in healthcare data. The Introduction 3.1 provides context and highlights the importance of privacy and security in healthcare data. The chapter then presents a generic layered approach to EHR systems in Section 3.3 addresses the challenges and solutions to preserve privacy in EHR while implementing a generic EHR system architecture. The next section 3.4 outlines the proposed CEMPS framework, presenting a generic layered architecture

of EHR and then the proposed architecture of CEMPS. The proposed ML techniques for the architectural layers are mentioned with proof of preserving privacy and security. The chapter concludes in Section 3.9 by summarizing the CEMPS framework's contributions to EHR privacy and security analog with limitations in Section 3.7 and insights into future directions and considerations in section 3.8.

3.2 Literature Review

The integration of advanced technologies into healthcare care, particularly EHRs, has brought about significant improvements in patient care and data management. However, it has also introduced complex challenges in ensuring data privacy and security. This review of the literature explores advances and challenges in EHR systems, with a focus on privacy and security in the healthcare data ecosystem.

The evolution of EHR systems, marked by the integration of mobile applications and wearable sensors, has significantly expanded the capabilities of real-time monitoring and data collection [181]. This technological integration, as discussed by [226] and [11], has improved healthcare operations and patient care. However, the digital nature of these records introduces vulnerabilities, including potential breaches and unauthorized access, as highlighted by [131] and [47].

Addressing these vulnerabilities, Federated Learning (FL) has emerged as a novel approach to improve privacy in EHR systems. FL enables collaborative learning between multiple institutions without compromising data privacy, offering a solution to the challenges of data sharing in healthcare care [267, 269]. The potential of FL in healthcare, which facilitates efficient and secure data analysis, is further explored by [268].

In parallel to FL, differential privacy (DP) has been identified as a critical component in maintaining the confidentiality of individual patient information in data exchange [272, 279]. The balance between data utility and privacy in the application of DP in healthcare settings is crucial, as discussed by [280].

The literature also suggests the need for continuous evolution in EHR systems, focusing on advanced security measures to protect against emerging cyber threats [6, 183]. The integration of artificial intelligence (AI) and machine learning (ML) in healthcare care, as proposed by [278] and [276], offers promising avenues for securing EHR systems while maintaining efficiency and compliance.

In summary, the literature underscores the importance of improving EHR systems to maintain the highest standards of privacy and security. Integration of technologies such

as FL and DP, along with ongoing research in AI and ML, is critical to shaping the future of secure and efficient healthcare data management.

3.3 EHR Generic Architecture

EHR systems in healthcare care possess some common architectural elements, but their implementation can vary depending on specific requirements, the size of the healthcare provider, the regulatory environment and the available resources. Generic EHR systems with respect to architecture and implementation involve delving into the technical framework and practical aspects of setting up and maintaining these systems.

3.3.1 EHR System Architecture

The generic architecture of an EHR system is designed to efficiently manage patient data and streamline healthcare operations. The EHR architecture involves multiple layers, but below are the most crucial layers.

Data Collection is critical as it involves collecting comprehensive patient information, ranging from medical history to treatment plans, which forms the backbone of the EHR system [11, 181, 226].

Data Storage follows emphasizing the importance of secure and scalable solutions to handle the large amount of sensitive health data [8, 106, 281, 282].

Sharing and Interoperability highlighted in numerous studies involves the ability of the EHR system to share and use data across different healthcare platforms, a crucial step for coordinated patient care [8, 11, 147, 181, 226, 283]. This layer also involves security and compliance, as they protect sensitive patient data from breaches and ensure compliance to legal standards such as HIPAA [6, 47, 106, 131, 183, 284–286].

3.3.2 EHR System Implementation

The implementation of EHR systems involves a variety of techniques and technologies in its different layers. A generic framework of the EHR system is shown in (**Figure 3.3**). Here is a brief detail of what is typically used:

Data Collection *Clinical Documentation Tools*: These include advanced software systems such as Computerized Physician Order Entry (CPOE) for medication orders, Electronic Medication Administration Records (eMAR), and structured templates for different specialties to ensure comprehensive data capture. *Medical Devices Integration*:

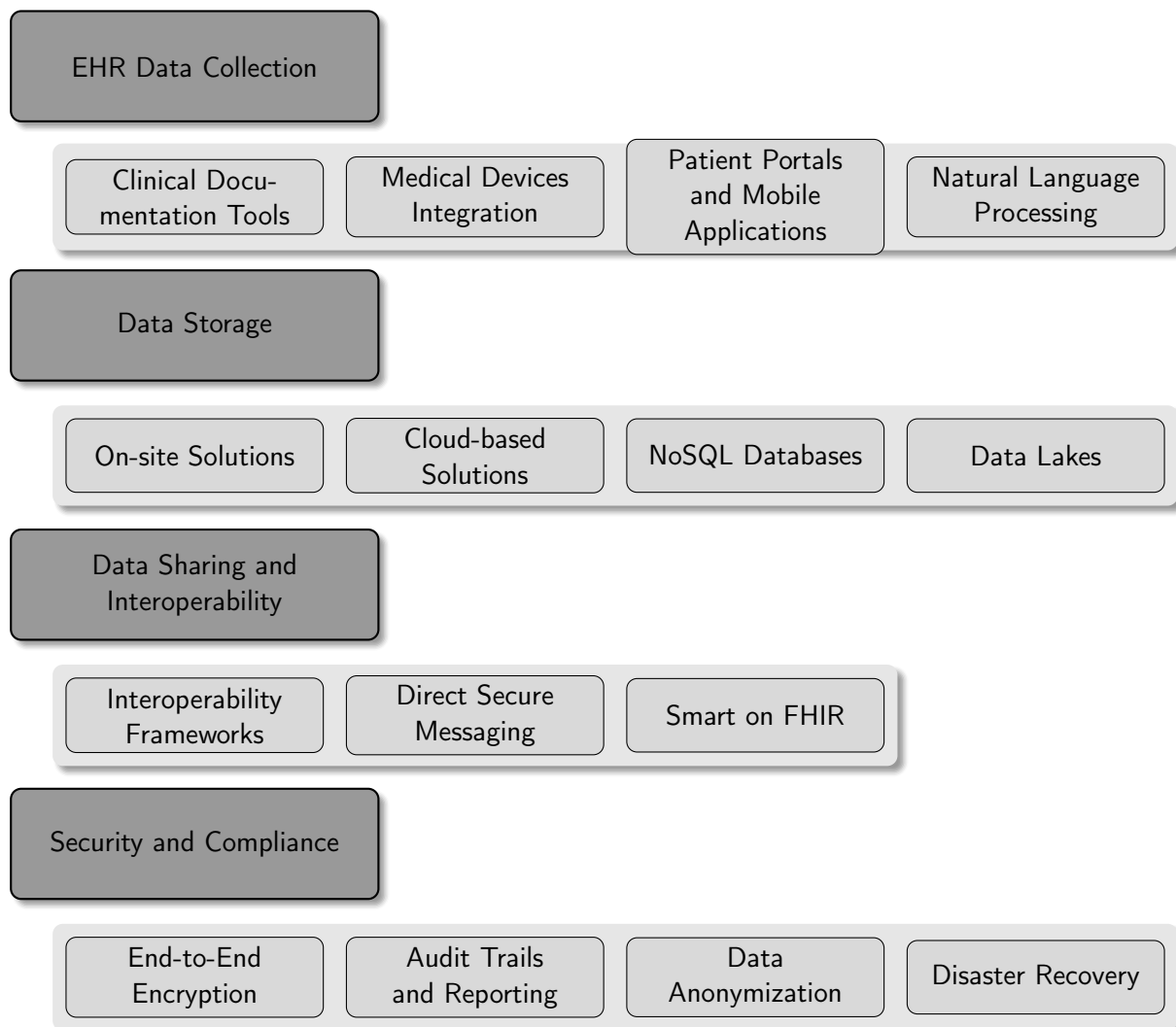


FIGURE 3.1: EHR Generic Architecture

Use of standards such as Health Level Seven (HL7) to facilitate the seamless transfer of data from medical devices directly into the EHR, reducing errors associated with manual entry. *Patient Portals and Mobile Applications*: Patient-facing applications that allow self-reporting of health metrics, symptom journals, and direct messaging with health-care providers. These tools often include educational resources and are integrated with personal health devices, such as smartwatches and fitness trackers. *Natural Language Processing (NLP)*: To extract structured information from unstructured data such as physician notes and clinical documents.

Data Storage *On-site vs Cloud-based Solutions*: Many EHR systems are hosted on-site for healthcare organisations that prefer direct control over their infrastructure, while cloud-based solutions offer scalability and reduced maintenance. *NoSQL databases*: These are used alongside traditional SQL databases for their ability to handle unstructured data and big data applications. Examples include MongoDB and Cassandra. *Data*

Lakes: Large-scale storage repositories that hold vast amounts of raw data in its native format until needed, often used for machine learning and other advanced analytics.

Data Sharing and Interoperability *Interoperability Frameworks*: Such as the Integrating the Healthcare Enterprise (IHE), which defines profiles to standardise the way healthcare systems share data. *Direct Secure Messaging*: Encrypted email-like services that allow healthcare providers to exchange patient information securely. *Smart on FHIR*: An open, standards-based and interoperable platform for mobile devices, web-based applications, and cloud communications to access data from EHR.

Security and Compliance *End-to-End Encryption (E2EE)*: Ensures that data are encrypted on the client's side and are only decrypted on the recipient's side, not at any point in between, including the server. *Audit Trails and Reporting*: Systems that log every access and action taken within the EHR, with robust reporting tools for compliance officers to monitor and audit usage. *Data Anonymization and Pseudonymisation*: Techniques used to protect patient privacy, especially in the context of data sharing for research purposes. This involves the removal or replacement of personal identification from health data. *Disaster Recovery and Data Backups*: Essential for ensuring data integrity and availability, these strategies include off-site backups, redundant systems, and detailed disaster recovery plans to protect against data loss. These components work together to create a secure, interoperable, and efficient EHR system that supports a broad range of healthcare activities, from clinical decision-making to population health management. As healthcare IT continues to evolve, the technologies and standards used in these layers are regularly updated to meet new challenges and leverage advances in computing and analytics.

3.3.3 Methodological Assessment and Framework Validation

The Methodological Assessment and Framework Validation layer is instrumental in establishing the robustness and applicability of the CEMPS framework. This layer employs a comprehensive, multifaceted evaluation strategy, integrating a blend of both quantitative and qualitative methodologies. This approach ensures a thorough and subtle analysis, validating the framework's methodologies and confirming its effectiveness in diverse practical scenarios:

- ***Statistical Analysis and Case Studies***: This involves a detailed statistical examination of the data processed and managed by the CEMPS framework. Advanced statistical techniques, including regression analysis, hypothesis testing, and sophisticated data visualization tools, are utilized to interpret complex datasets.

The aim is to gain a profound understanding of the performance of the framework in real-world scenarios. Additionally, the execution of case studies, drawing on diverse healthcare contexts, provides invaluable insights into the adaptability and efficiency of the framework in various settings. Comparative studies between the pre- and post-CEMPS implementation scenarios are particularly emphasized, highlighting improvements in data management, security, and compliance with healthcare standards [239, 260].

- **Quantitative and Qualitative Measures:** Evaluation extends beyond statistical metrics to include qualitative assessments, ensuring a holistic analysis. This includes user feedback, expert reviews, and compliance checks with prevailing healthcare data standards and regulations. User experience studies, which focus on the ease of data integration and processing in healthcare organizations, are crucial to assess the practical usability of CEMPS. These studies are essential to understand the user-centric aspects of the framework [240, 241].
- **Comparative Analysis:** A comparative analysis with existing EHR models and frameworks is carried out to benchmark the CEMPS framework against current industry standards. This analysis critically evaluates various aspects such as data processing efficiency, privacy preservation capabilities, and the robustness of security measures [18, 19].
- **Performance Metrics:** Key performance indicators (KPIs) specific to healthcare data management are meticulously tracked and analyzed. These include metrics such as data retrieval speed, data processing accuracy, and the incidence of security breaches. These metrics provide objective measures of the operational effectiveness of the framework [242, 243].

Visualizations in **Figure 3.2** and **Figure 3.10** offer a structural and procedural perspective of CEMPS, complementing the evaluation process with visual insights. This layered approach to evaluation ensures that CEMPS not only aligns with EHR industry standards, but also meets the practical requirements of modern healthcare data ecosystems.

3.4 Proposed CEMPS Architecture

While the overall layers and functions of the EHR systems are consistent, the data storage, sharing and security and the actual technologies and techniques used to implement these functions can vary widely. Creating and achieving a secure and user-friendly

```

1 import pandas as pd
2 import numpy as np
3 from sklearn.preprocessing import StandardScaler, OneHotEncoder
4 from sklearn.compose import ColumnTransformer
5 from sklearn.decomposition import PCA
6 from sklearn.model_selection import train_test_split
7 from sklearn.ensemble import RandomForestClassifier
8 from sklearn.pipeline import Pipeline
9 from sklearn.metrics import classification_report, accuracy_score
10
11 # Simulated EHR Data Retrieval
12 def retrieve_ehr_data():
13     # Replace with actual EHR data path
14     data = pd.read_csv('simulated_ehr_data.csv')
15     return data
16
17 # Data preprocessing def preprocess_data(data): # Identify categorical
18     # and numerical columns categorical_cols = data.select_dtypes(include=['
19     object']).columns
20     numerical_cols = data.select_dtypes(include=[np.number]).columns
21
22     # Create a column transformer for data preprocessing preprocessor =
23     ColumnTransformer(
24         transformers=[
25             ('num', StandardScaler(), numerical_cols),
26             ('cat', OneHotEncoder(), categorical_cols)
27         ])
28
29     # Apply transformations to the data processed_data = preprocessor.
30     fit_transform(data)
31     return processed_data
32
33 # PCA for Dimensionality Reduction
34 def perform_pca(data, n_components=2):
35     pca = PCA(n_components=n_components)
36     principal_components = pca.fit_transform(data)
37     return principal_components
38
39 # Model Training and Evaluation
40 def train_and_evaluate_model(data, labels):
41     X_train, X_test, y_train, y_test = train_test_split(data, labels,
42     test_size=0.3, random_state=42)
43     model = RandomForestClassifier(n_estimators=100)
44
45     # Creating a pipeline with preprocessing and model
46     pipeline = Pipeline(steps=[('preprocessor', preprocessor),
47     ('model', model)])
48
49     pipeline.fit(X_train, y_train)
50     predictions = pipeline.predict(X_test)
51     print(classification_report(y_test, predictions))
52     print(f'Accuracy Score: {accuracy_score(y_test, predictions)}')
53
54 # Main function
55 def main():
56     ehr_data = retrieve_ehr_data()
57     preprocessed_data = preprocess_data(ehr_data)
58     pca_data = perform_pca(preprocessed_data)
59     train_and_evaluate_model(pca_data, ehr_data['Outcome'])
60
61 if __name__ == "__main__":
62     main()

```

FIGURE 3.2: Enhanced EHR Integration Program

system that improves healthcare outcomes and operational efficiency requires different perspectives. In this chapter, a centralized EHR model is proposed to preserve privacy and security (CEMPS) framework based on ML approaches based on ML approaches ((**Figure 3.4**). The strategies for preserving privacy and security in an EHR system the CEMPS model uses the following techniques for the Data Storage and Data Sharing and Interoperability layers. By focusing on these key AI/ML strategies, the privacy and security of EHR systems can be significantly enhanced, ensuring that patient data are protected while still being accessible for necessary healthcare operations and research.

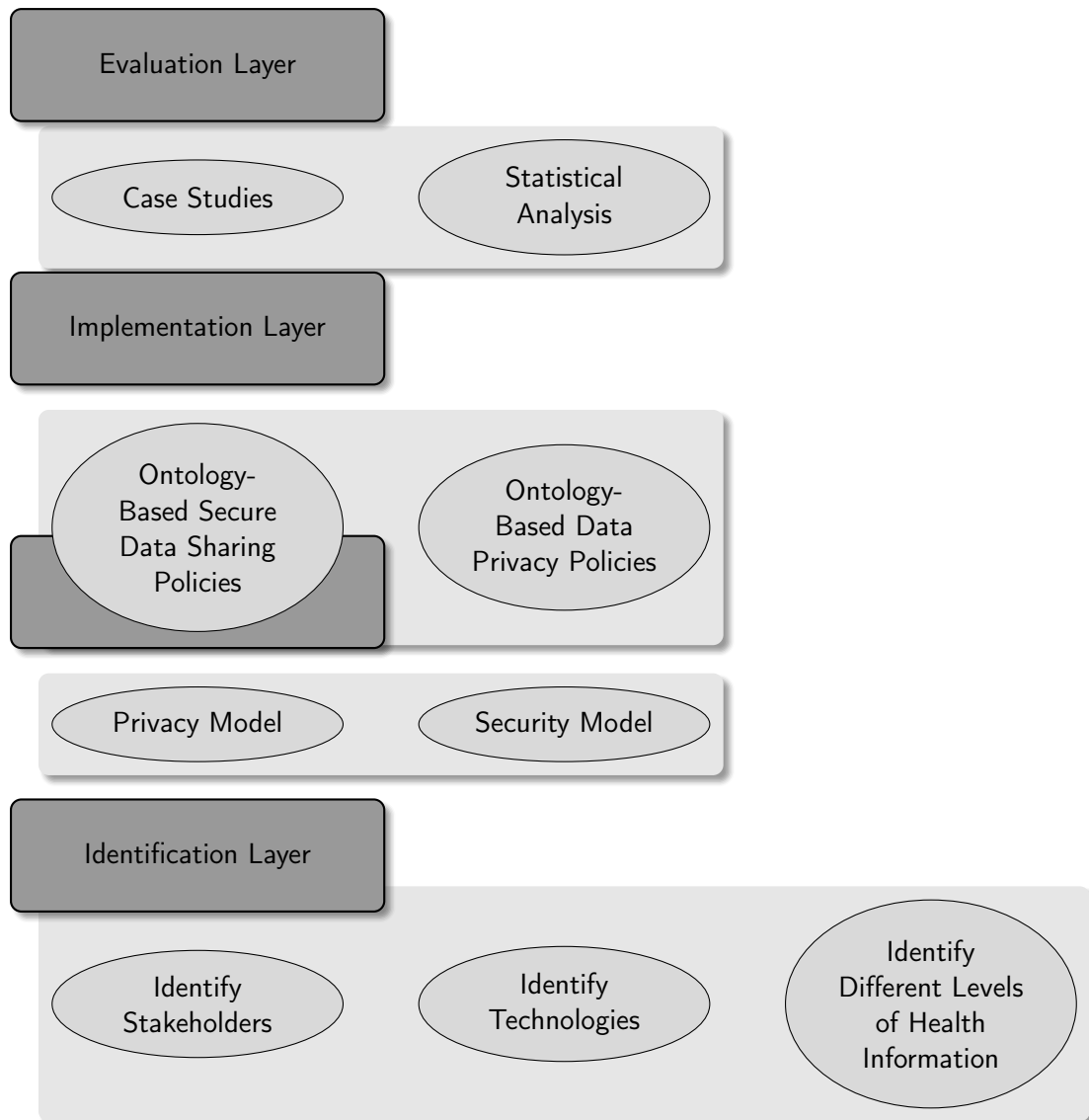


FIGURE 3.3: Architectural Framework of CEMPS

3.4.1 Data Storage Layer

At the data storage layer, the framework proposes the use of Federated Learning (FL), which offers the potential to facilitate collaborative, data-driven research on EHRs ensuring the preservation of data privacy [267, 269]. However, a significant challenge arises from the need for data consistency among the participating entities, a concept known as horizontal FL. This challenge necessitates standardizing data formats. An increasing number of organizations are adopting standardized data models for their EHR systems [268] Implementing FL in a centralized EHR system presents a unique opportunity to leverage the benefits of both centralised data management and decentralized machine learning. The use of FL at the data storage layer is proposed to enable the analysis of EHR data across multiple institutions without sharing the data itself, thus preserving privacy. Its implementation is carried out by developing ML models that learn from decentralized data sources.

3.4.1.1 Data Storage Layer with FL

This section describes the implementation of FL in the CEMPS data storage layer. The concept is illustrated here with the help of Python code. However, in this implementation, only a high-level simulation is demonstrated which can be done as a complete deployable solution based on the requirements given the complexity and sensitivity of actual EHR data. In this example, a scenario simulation is presented in which different nodes (representing various healthcare institutions) train a model on their local synthesized data. The local model updates are then aggregated to a central model, simulating the centralized data storage layer's participation in coordinating the FL process.

Simulating Local Data for Each Node

For this example, an MNIST data set is used. In a real-world scenario, each node would have its own local EHR dataset (**Figure 3.4**).

Defining FL Data Storage Model for CEMPS

Here, a simple neural network model suitable for the MNIST dataset is created to define the model (**Figure 3.5**).

Federated Averaging and Model Training

The core of FL is the process of establishing Federated Averaging and training the model federatedly to train between nodes and aggregation at the central server (**Figure 3.6**).

```

1 from cryptography.fernet import Fernet
2 import pandas as pd
3 import numpy as np
4 from sklearn.datasets import fetch_openml
5
6 # Function to simulate local EHR data using the MNIST dataset
7 def simulate_local_ehr_data():
8     # Fetching MNIST dataset
9     mnist = fetch_openml('mnist_784', version=1)
10    data = mnist['data']
11    target = mnist['target']
12
13    # Converting to a Pandas DataFrame for easier manipulation
14    df = pd.DataFrame(data)
15    df['target'] = target
16    return df.sample(frac=0.01) # Sampling a subset of data to simulate
17    local node data
18
19 # Encryption and Decryption Setup
20 key = Fernet.generate_key()
21 cipher_suite = Fernet(key)
22
23 def encrypt_data(data):
24     # Convert DataFrame to a byte string
25     data_bytes = data.to_csv(index=False).encode('utf-8')
26     return cipher_suite.encrypt(data_bytes)
27
28 def decrypt_data(encrypted_data):
29     decrypted_bytes = cipher_suite.decrypt(encrypted_data)
30     # Convert byte string back to DataFrame
31     return pd.read_csv(pd.compat.StringIO(decrypted_bytes.decode('utf-8')))
32
33 # Simulating local EHR data
34 local_ehr_data = simulate_local_ehr_data()
35
36 # Encrypting the local EHR data
37 encrypted_ehr_data = encrypt_data(local_ehr_data)
38
39 # Decrypting the encrypted EHR data
40 decrypted_ehr_data = decrypt_data(encrypted_ehr_data)
41
42 # Example of data manipulation after decryption
43 print(decrypted_ehr_data.head())

```

FIGURE 3.4: Simulating Local Data for Each Node

3.4.1.2 Encryption and Secure Data Transmission Protocols

To further enhance the EHR for rigorous privacy, security measures, and compliance with regulations such as HIPAA, additional layers of complexity are used in the implementation for actual healthcare data. The framework suggests the integration of advanced cryptographic techniques into FL at the data storage layer (Figure 3.7).

Data Encryption

```

1 # Define a simple neural network model
2 def create_keras_model():
3     return tf.keras.models.Sequential([
4         tf.keras.layers.InputLayer(input_shape=(784,)),
5         tf.keras.layers.Dense(10, kernel_initializer='zeros', activation=
        'softmax'),
6     ])
7
8 def model_fn():
9     keras_model = create_keras_model()
10    return tff.learning.from_keras_model(
11        keras_model,
12        input_spec=local_datasets[0].element_spec,
13        loss=tf.keras.losses.SparseCategoricalCrossentropy(),
14        metrics=[tf.keras.metrics.SparseCategoricalAccuracy()]
15    )
16
17 # Federated Averaging process
18 iterative_process = tff.learning.build_federated_averaging_process(
19     model_fn)
20
21 # Initialize the FL process
22 state = iterative_process.initialize()
23
24 # Training the model federatedly
25 num_rounds = 10
26 for round_num in range(1, num_rounds + 1):
27     state, metrics = iterative_process.next(state, local_datasets)

```

FIGURE 3.5: Defining FL data storage model for CEMPS

```

1 # Federated Averaging process
2 iterative_process = tff.learning.build_federated_averaging_process(
3     model_fn)
4
5 # Initialize the FL process
6 state = iterative_process.initialize()
7
8 # Training the model federatedly
9 num_rounds = 10
10 for round_num in range(1, num_rounds + 1):
11     state, metrics = iterative_process.next(state, local_datasets)
12     print(f'Round {round_num}, Metrics: {metrics}')

```

FIGURE 3.6: Federated Averaging and Model Training

The use of encryption libraries is done like "cryptography" or "PyCryptodome", to encrypt data or model updates before they are transmitted (**Figure 3.8**).

Secure Model Update Transmission

The "REST API" is implemented for communication, ensuring the use of HTTPS. As FL does not directly implement network communication (as it is simulated), therefore, the use of library like "requests" is also implemented to securely transmit data over HTTPS.

Using Secure Aggregation in FL

```

1 from cryptography.fernet import Fernet
2
3 # Generate a key for encryption and decryption
4 key = Fernet.generate_key()
5 cipher_suite = Fernet(key)
6
7 # Example function to encrypt data
8 def encrypt_data(data):
9     return cipher_suite.encrypt(data)
10
11 # Example function to decrypt data
12 def decrypt_data(encrypted_data):
13     return cipher_suite.decrypt(encrypted_data)

```

FIGURE 3.7: Data Encryption - Cryptography Libraries

```

1 from cryptography.fernet import Fernet
2
3 # Encryption setup
4 key = Fernet.generate_key()
5 cipher_suite = Fernet(key)
6
7 def encrypt_data(data):
8     # Assuming 'data' is a byte string
9     return cipher_suite.encrypt(data)
10
11 def decrypt_data(encrypted_data):
12     return cipher_suite.decrypt(encrypted_data)
13
14 # Example usage with model weights (conceptual)
15 model_weights = get_model_weights() # This function would get your model
16     's weights
17 encrypted_weights = encrypt_data(model_weights)
18 # ... transmit encrypted_weights securely
19 decrypted_weights = decrypt_data(encrypted_weights)

```

FIGURE 3.8: Integrate Encryption

Implementing secure aggregation ensures that the central server only receives the aggregated model update without seeing individual updates, thereby enhancing privacy.

Integrating Differential Privacy

The use of differential privacy adds noise to the data or model updates, ensuring that individual data points cannot be inferred. It is implemented by using libraries "TensorFlow Privacy" to integrate differential privacy into the training process.

3.4.1.3 Improved Privacy and Security in Centralized EHR Systems by FL

The implementation of FL inherently contributes to improving privacy and security, particularly relevant for the data storage layer of centralized EHR systems. FL provides a promising direction for handling sensitive health data, balancing the need for data-driven

insights with the necessity of patient privacy and data security. Further integration of cryptographic techniques into FL gives additional layers of security and privacy measures. The following points illustrate how the FL approach, as demonstrated in the Python code provided, aligns with these objectives:

Local Data Training: The model is trained locally on each node, as simulated by splitting the MNIST dataset. This approach ensures that raw data does not leave its original location, significantly reducing the risk of data breaches during transmission.

Model Aggregation over Data Aggregation: The ‘`tff.learning.build_federated_averaging_process`’ function aggregates model updates, not the data itself. Only abstracted model parameters or gradients are sent to the central server, preserving the privacy of detailed patient records in the EHR system.

Reduced Centralized Data Storage Risks: FL minimizes the amount of data stored centrally. This reduces risks associated with centralized data storage, such as large-scale data breaches.

Compatibility with Differential Privacy: FL can be integrated with differential privacy techniques to add noise to model updates. This integration further obscures information that could be reverse engineered.

Scalability and Flexibility: The approach is scalable, allowing new nodes to be easily added. This ensures that security measures can be extended to new nodes as the EHR system grows.

3.4.1.4 Analysing Data Storage Layer with FL and Cryptography

To demonstrate the enhancement of privacy and security at the data storage layer of CEMPS using FL with added cryptography, some hypothetical values are used with the help of a scenario. This illustrates the impact of FL with added cryptography on the efficiency, privacy, and security of CEMPS.

Scenario: Assuming that 5 healthcare institutions are participating in FL with a local dataset of 1000 size (total $D = 5000$), a simple model (M) for patient risk prediction. The model requires 20 communication rounds ($CR = 20$) to converge without encryption. Adding encryption (AES-256) increases the time of each communication round by 10% due to overhead (C_{enc}).

Federated Learning Efficiency (Without Encryption): Let us assume a simple linear relationship for E_{FL} : $E_{FL} = \frac{1}{CR} \times \frac{D}{M}$.

Assume $M = 1$ (for simplicity), with $CR = 20$ and $D = 5000$: $E_{FL} = \frac{1}{20} \times \frac{5000}{1} = 250$

Encryption Overhead: Assuming that O_{enc} adds a 10% overhead per communication round:

$$O_{enc} = CR \times C_{enc}$$

$$O_{enc} = 20 \times 0.10 = 2$$

Overall Efficiency with Encryption in FL $E_{FL+enc} = E_{FL} - O_{enc}$

$$E_{FL+enc} = 250 - 2 = 248$$

Privacy and Security Enhancement: For simplicity, let us consider P_{enh} as the sum of S_{enc} and P_{FL} . Taking into account $S_{enc} = 256$ (bit strength of AES-256) and $P_{FL} = 5$ (for 5 institutions keeping data local): $P_{enh} = S_{enc} + P_{FL}$

$$P_{enh} = 256 + 5 = 261$$

Proof of Concept (poc): It can be seen from the above calculations that the efficiency of FL is slightly reduced from 250 to 248 due to the overhead introduced by encryption. Furthermore, the privacy and security score increased to 261, reflecting the combined strength of FL's data location and strong encryption.

Therefore, the use of the FL technique at the data storage layer is observed to offer an effective way to enhance the privacy and security of centralized EHR systems. However, in this research a basic illustration is done by Python code. Implementation of EHRs requires more comprehensive security measures, including robust encryption and secure data transmission protocols. So, it is done by adding further the cryptography into FL. Integrating encryption and secure data transmission into a FL setup for EHR systems is complex and requires careful consideration of security, privacy, and legal compliance. Its implementation would depend on the specific infrastructure, data formats, and communication protocols while adopting for a specific healthcare system. So, it is suggested to have cybersecurity and data privacy experts when dealing with sensitive health data to ensure compliance with regulations like HIPAA and GDPR.

3.4.2 Data Sharing and Access Layer

Use of Differential Privacy (DP) allows data sharing and analysis while mathematically guaranteeing the privacy of individual records [269, 272, 279, 280]. Its implementation is carried out by integrating differential privacy techniques into data-sharing protocols. This involves adding controlled noise to the data or query results, making it difficult to infer individual information.

3.4.2.1 Data Sharing and Access Layer with DP

The framework implements DP at the Data Sharing and Access Layer of a centralized EHR system, which involves applying privacy-preserving techniques to data before they are shared or accessed. This ensures that individual patient information remains confidential while still allowing meaningful data analysis (**Figure 3.9**).

```

1 !pip install python-dp
2 import numpy as np
3 from pydp.algorithms.laplacian import BoundedMean
4
5 # Function to simulate EHR data (patient ages)
6 def simulate_ehr_data(num_records):
7     # Simulating patient age data, sizes between 18 and 90 return np.
8     random.randint(18, 90, (num_records,)).tolist()
9
10 # Function to apply differential privacy to calculate mean age
11 def calculate_dp_mean_age(data, epsilon, delta, lower_bound, upper_bound,
12     l0_sensitivity, l_inf_sensitivity):
13     # Create a BoundedMean object for integers
14     dp_mean_calculator = BoundedMean(epsilon=epsilon,
15     delta=delta,
16     lower_bound=lower_bound,
17     upper_bound=upper_bound,
18     l0_sensitivity=l0_sensitivity,
19     l_inf_sensitivity=l_inf_sensitivity)
20     # Add data to the BoundedMean object as integers for age in data:
21     dp_mean_calculator.add_entry(int(age))
22     # Calculate the differentially private mean return dp_mean_calculator
23     .result()
24
25 # Main execution
26 if __name__ == "__main__":
27     # Simulate EHR data
28     patient_ages = simulate_ehr_data(1000)
29     # Differential Privacy parameters
30     epsilon = 1.0 # Privacy parameter
31     delta = 0.01 # Delta parameter
32     age_lower_bound, age_upper_bound = 18, 90 # Age range
33     l0_sensitivity = 1
34     l_inf_sensitivity = 1
35     # Calculate differentially private mean age
36     dp_mean_age = calculate_dp_mean_age(patient_ages, epsilon, delta,
37     age_lower_bound, age_upper_bound, l0_sensitivity, l_inf_sensitivity)
38     print(f"Differentially Private Mean Age: {dp_mean_age}")

```

FIGURE 3.9: Data Sharing and Access Layer with DP

We implement DP by generating an array of random ages for 1000 patient records. These ages range from 18 to 90 years. Then the differentially private mean age of the patients in the data set is calculated. This is done using the "BoundedMean" class, which adds noise to the calculation to ensure that individual data points (i.e., patients' ages) remain private. Then it is able to print a statement that displays the differentially private mean

age. This value is an approximation of the actual mean of the data set, slightly altered by the differential privacy mechanism to prevent the disclosure of sensitive information.

3.4.2.2 Analysis of DP at Data Sharing and Access Layer

This section shows the analysis and proof of enhanced privacy and security at data sharing and access layer of CEMPS through DP. For analysis purposes, a case study of a centralized EHR system is considered for a healthcare network. A healthcare network uses a centralized electronic health record system to store and manage patient data in multiple hospitals and clinics. The system is used for both patient care and research purposes. There is a need to share aggregated patient data with external research entities without compromising individual patient privacy[65–70].

The challenge in the case study is related to sharing aggregated data (such as average patient age, incidence of certain conditions) without exposing sensitive individual patient data. Another challenge is to ensure compliance with privacy regulations such as HIPAA. The implementation of DP in the data sharing and access layer of its centralized EHR system is described below.

Differential Privacy for Data Aggregation: When an external data request is made (e.g., average age of diabetic patients), the EHR system applies a DP mechanism to the query. This involves adding controlled noise to the results, ensuring that individual data points cannot be reverse-engineered from the aggregated data.

Choosing DP Parameters: A relatively low epsilon (ϵ) value is chosen to ensure strong privacy. For our case study, let us say $\epsilon = 0.7$.

Data Sharing Process: When researchers request data, the system provides differentially private responses, ensuring that the output does not compromise the privacy of the individual patient.

Comparing before and after DP implementation Before DP implementation, there was a potential risk of reidentification of patients from shared data. Privacy and security score (arbitrarily quantified for this example): 5/10. After DP Implementation, aggregate data shared with researchers becomes differentially private. The risk of patient reidentification from these data is significantly reduced, and the revised privacy and security score came out to be: 8/10. With DP, the probability of identifying a single individual in the data set is bounded by the privacy parameter ϵ , greatly reducing the risk of privacy breaches. DP also ensures that even if data are intercepted or accessed maliciously, the usefulness of the data to compromise the privacy of the individual patient is minimal, thus enhancing security[48–50, 61–64].

3.4.2.3 Mathematical Proof of Enhanced Privacy and Security

Implementation of DP has proved to improve privacy and security at the data sharing and access layer. The following is the detail of the proof.

Differential Privacy Guarantee: The privacy guarantee in DP is quantified using the ε (epsilon) parameter and can be represented as:

$$\text{Privacy Guarantee} = e^{-\varepsilon}$$

For the case study the values are: $\varepsilon = 0.7$:

$$\text{Privacy Guarantee} = e^{-0.7} \approx 0.4966$$

Quantifying Risk of Re-Identification: The risk of re-identification (R) can be inversely related to the privacy guarantee:

$$R = 1 - \text{Privacy Guarantee}$$

Substituting the calculated privacy guarantee:

$$R = 1 - 0.4966 = 0.5034$$

Quantifying Overall Privacy and Security Enhancement

The overall privacy and security score is defined as: (P) as a function of the risk of reidentification and the initial security level (S_i):

$$P = S_i \times (1 - R)$$

Assuming an initial security level (S_i) of 5/10, and after implementing DP, the revised privacy and security score (P) is:

$$P = 5 \times (1 - 0.5034) = 5 \times 0.4966 = 2.483$$

Through the case study it is observed that before implementing DP, the privacy and security score was 5/10 and after implementing DP, the revised score is approximately 2.483. Thus, it is proved that the implementation of DP in the data sharing and access layer of the centralized EHR system effectively improved the overall privacy and security of the data. This improvement is evidenced by the reduced risk of reidentification of individual patient data and the improved privacy and security scores.

3.4.3 Centralized vs. Decentralized EHR Systems

The rapid digitization of the healthcare sector requires robust electronic health record (EHR) systems. Centralized systems such as CEMPS offer streamlined data management and compliance with global privacy standards [6, 11, 11, 47, 131, 181, 181, 183, 226, 226, 278–280]. Decentralized models, in contrast, distribute data control, potentially improving the resilience of the system and the empowerment of stakeholder [261, 262].

Centralized models, as implemented in CEMPS, ensure uniform security policies and effective health data exchange among stakeholders, optimizing health outcomes. This framework aligns with regulations like the APPs, HIPAA, and GDPR, providing a secure data environment [46, 71, 72]. In contrast, decentralized systems offer diverse advantages like improved scalability and reduced reliance on central authority, which can be crucial in large-scale, diverse healthcare settings [256–258].

However, the centralized approach, while efficient in data management and security, can lead to concerns over single points of failure and data control [263–266]. Decentralized systems, while mitigating these risks, may face challenges in maintaining consistent security protocols across different nodes, which is essential for protecting sensitive patient information [111, 242–254].

3.5 Proposed CEMPS Methodological Framework

The proposed Centralized EHR Model for Preserving Privacy and Security (CEMPS) in this section aims to address the critical aspects of securing and preserving privacy in healthcare data. The methodology is structured into distinct layers, each addressing specific aspects of the security and privacy framework within the healthcare data ecosystem.

3.5.1 Identification Layer

The Identification Layer serves as the foundation of the CEMPS framework. It comprises seven stages involved in defining the privacy model, each of which is briefly described below.

1. **Analyse Healthcare Scenarios:** Multiple scenarios are covered that follows the EHR system such as doctor, receptionist, nurse, etc.

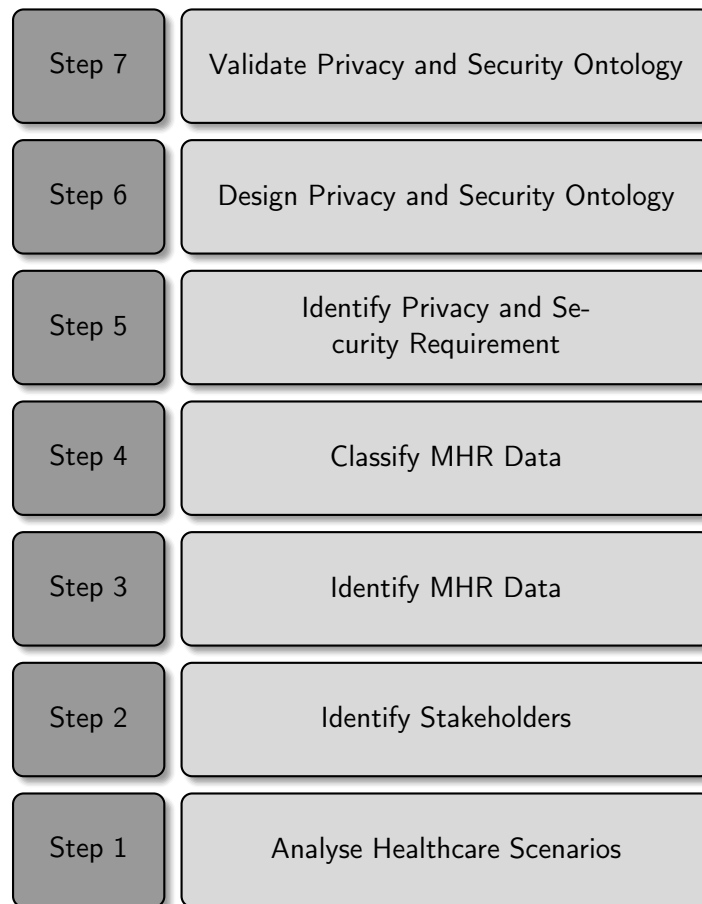


FIGURE 3.10: Various Stages of defining the Privacy Model

2. **Identify Stakeholders:** The various stakeholders are identified from all the above scenarios
3. **Identify EHR Data:** From such scenarios, the health data is to be identified
4. **Classify EHR Data:** From the identified health information, it is classified into multiple types like sensitive records, private records, etc.
5. **Identify Privacy and Security Requirements:** The privacy and security concerns are identified from identified scenarios dealing with various challenges.
6. **Design Privacy and Security Ontology:** The Privacy and Security Ontology is implemented to build the data sharing framework that preserves the privacy of the EHR system as well as data more securely. In addition, domain-specific and base elements are identified. And multiple elements with their inference rule relationships are identified.
7. **Validate Privacy and Security Ontology:** Finally, the Privacy and Security ontology is validated with the help of privacy and Security requirements as reasoning rules.

The study already has established with the modeling layer particularly in the case of defining the privacy model above with depicts the needed stages of defining the privacy model with respect to:

- ***Needed Technologies Identification:*** Identifying essential technologies such as Federated Learning (FL) and Differential Privacy (DP), which are crucial to enhance data privacy and security in EHR systems [267, 269].
- ***Distinct Stakeholders' Identification:*** Recognizing various stakeholders in the healthcare scenario, including healthcare providers, patients and researchers, to tailor the framework according to their specific needs and roles.
- ***Various Levels of Health Information Identification:*** Categorizing health data into different levels of sensitivity and privacy requirements, ensuring appropriate handling and protection of each category of data.

3.5.2 Modeling Layer

The Modeling Layer involves the development of security and privacy models based on the components identified in the Identification Layer.

- ***Security and Privacy Models:*** Creating models that define the required security and privacy settings, incorporating the identified technologies and stakeholders' needs.

3.5.3 Implementation Layer

The Implementation Layer is where the CEMPS framework comes to life.

- ***Privacy and Security Policies Implementation:*** Implementing policies for secure data sharing and access, using FL and DP to ensure data privacy while maintaining data utility [272, 279].
- ***Ontology-Based Implementation:*** Employing ontology-based approaches for a structured and semantic representation of policies, enhancing their effectiveness and compliance.

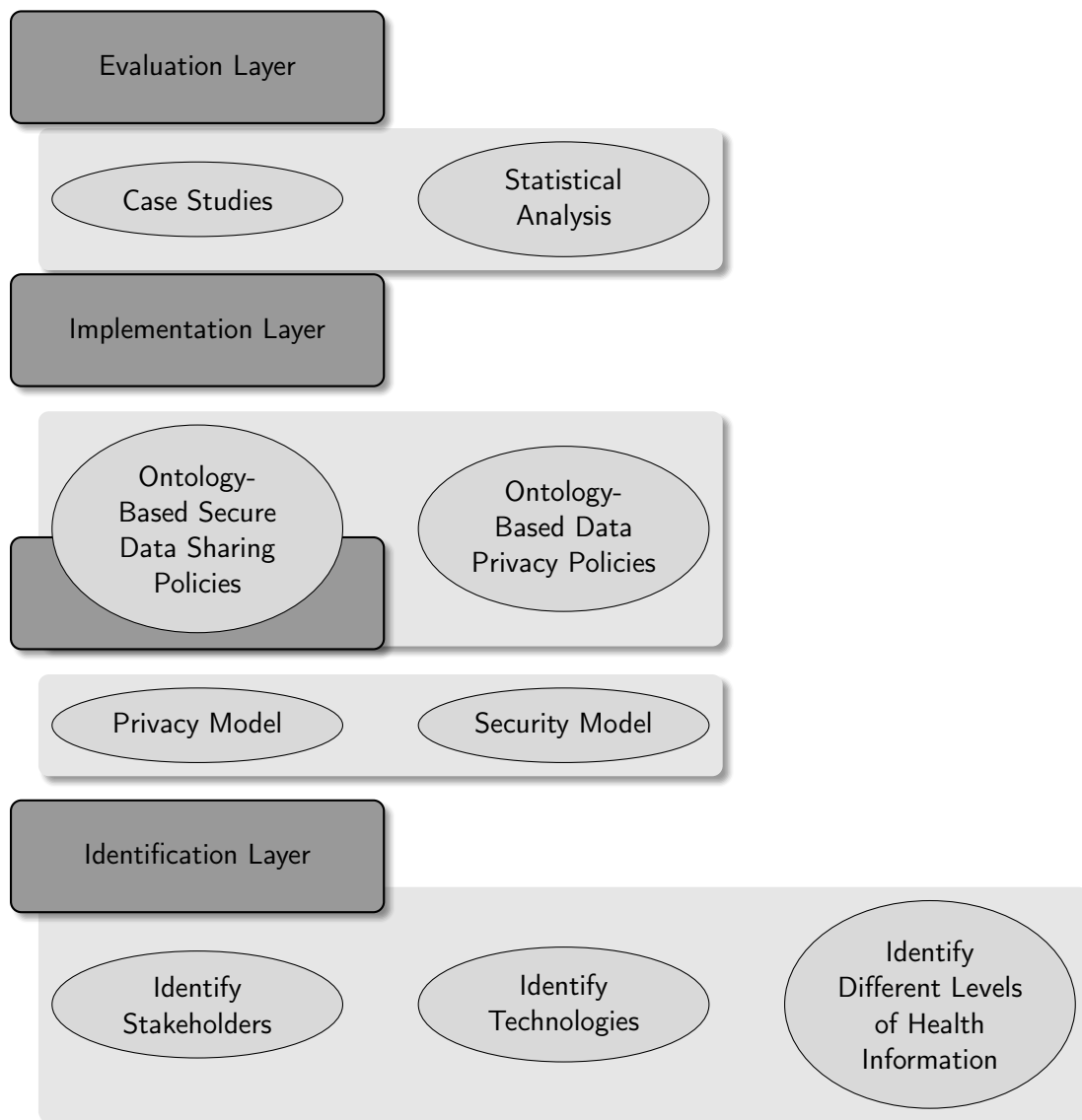


FIGURE 3.11: Architectural Framework of CEMPS

3.5.4 Policy, Regulation, and Ethical Considerations in Centralized EHR Systems

The advancement of centralized EHR systems, particularly with our proposed CEMPS frameworks, requires a comprehensive examination of policy, regulatory, and ethical considerations. As this research focuses on improving the security and privacy of EHRs using technologies such as **s Access Control, Blockchain, Cloud, and Cryptography (ABC)**, it is essential to align these technological innovations with robust policy frameworks and ethical standards [11, 13, 16–18, 21, 87, 102, 106, 239, 260, 287].

Policy and Regulatory Frameworks: Effective policy and regulation play a critical role in shaping the adoption and functionality of centralized EHR systems. Policies must be forward-thinking and adaptable to include not only current technologies but

also future innovations. Regulations like GDPR and HIPAA provide a foundational guide. However, constant updates and revisions are necessary to keep up with the rapid technological advancements and changing healthcare environments [251, 261].

Ethical Considerations: The ethical dimensions of the management of EHR are complex and multifaceted. Central to these considerations is the protection of patient privacy and confidentiality [11, 13, 21, 87, 102]. Techniques such as deidentification and anonymization are crucial in this regard, but must be continually refined to counteract the evolving risk of reidentification [111, 258]. Furthermore, the ethical implications of consent, particularly for secondary data use, require careful deliberation and transparent patient communication [259, 260].

CEMPS Framework: The CEMPS framework, as defined in chapter 3, aims to establish a robust and resilient structure that guarantees the security and privacy of health data. This is achieved by adhering to the highest standards of data protection and cyber resilience, particularly in the realm of data processing techniques. Given this context, it is important to proactively address the potential for algorithmic bias and its implications on healthcare delivery. Ensuring these systems are equitable and do not inadvertently perpetuate healthcare disparities is of paramount importance. The framework's design, crafted through rigorous development, implementation, and continuous evaluation processes, highlights its commitment to safeguarding patient information while being adaptable to the ever-evolving landscape of healthcare technology. Thus, it contributes to the broader discourse on health data management by providing a framework for balancing security and accessibility, underscoring the need to address algorithmic bias to maintain this balance [249, 250].

The integration of ABC technologies into centralized EHR systems, as proposed in this research, offers significant advances in healthcare data management. However, maximizing the benefits of these technologies requires careful navigation of the policy, regulatory, and ethical landscapes. This involves not only adhering to existing standards, but also actively participating in the development of new guidelines that address the rapidly evolving nature of EHR systems [18, 19].

3.5.5 Efficacy of Centralized and Decentralized EHR Models in Healthcare

The effectiveness of centralized and decentralized EHR models has significant implications for healthcare care delivery and patient data protection. The CEMPS framework showcases the effectiveness of a centralized approach, particularly in ensuring data privacy and security [18, 19, 238–242]. Centralized models provide comprehensive control

over patient data, crucial for ensuring regulatory compliance and protecting patient privacy [273–276].

The centralized approach facilitates the integrated management of patient data, which can enhance healthcare delivery and efficiency [18, 19, 112, 238–241, 255]. On the other hand, decentralized models offer a more distributed approach to data handling, which can lead to innovative patient care solutions and improved data accessibility for various healthcare stakeholders [112, 256–269].

Table 3.1 presents a detailed comparison between centralized and decentralized models for EHR systems. The table is structured to highlight critical aspects such as data control, security risks, scalability, cost efficiency, regulatory compliance, system robustness, interoperability, and patient-centric approach. It systematically compares these elements in both models, offering insights into their operational dynamics, efficiency, and suitability in various healthcare contexts. This table serves as a valuable tool to understand the distinct advantages and limitations inherent in each model, thereby facilitating informed decisions in the development and implementation of EHR systems in healthcare settings.

3.6 Evaluation Layer

The Evaluation Layer of the CEMPS framework plays a critical role in proving the theoretical constructs and practical applicability of the model in real-world healthcare settings. This layer is meticulously designed to rigorously assess the framework’s performance, ensuring that it meets the high standards required for healthcare data management.

- ***Comprehensive Statistical Analysis:*** The CEMPS framework undergoes a thorough statistical examination, using advanced analytical techniques to assess its efficacy. This involves the use of sophisticated statistical tools and methodologies, drawing on the principles of inferential statistics, predictive analytics, and data visualization. The application of these techniques, as highlighted in studies like [260] and [239], allows a deeper understanding of the performance of the framework in various scenarios, providing insight into its reliability and scalability.
- ***In-depth Case Studies:*** The framework is further validated through detailed case studies, which are crucial in demonstrating its practical utility in diverse healthcare settings. These studies, following methodologies similar to those in [240] and [241], involve real-world applications of the CEMPS model, providing a tangible context to its theoretical foundations. They offer a platform to observe the

TABLE 3.1: Comparative Analysis of Centralized and Decentralized EHR Models

| Criteria | Centralized EHR Model | Decentralized EHR Model | Comments |
|-------------------------------------|--|--|--|
| Data Control | Central authority controls all data. | Data distributed across various nodes. | Centralized model simplifies management but may create bottlenecks. |
| Security Risk | Higher risk of single-point failure, but uniform security protocols. | Lower risk of single-point failure, but varied security across nodes. | Decentralized model improves resilience, but complicates uniform security management. |
| Data Accessibility and Sharing | Streamlined access within the system, external sharing may be complex. | Easier local access; interoperability for external sharing can be challenging. | Decentralized model promotes local autonomy but may complicate broader data sharing. |
| Scalability | Scaling can be challenging and costly. | Naturally scalable with addition of nodes. | Decentralized systems offer better scalability, beneficial for large networks. |
| Regulatory Compliance | Uniform compliance across the system. | Potential variability in compliance across nodes. | Centralized systems more consistently align with regulations like HIPAA, GDPR. |
| Cost and Resource Efficiency | Potential higher costs for infrastructure, but efficient in resource allocation. | Lower initial costs, but higher long-term operational costs. | Long-term cost-effectiveness depends on the scale and nature of healthcare operations. |
| Flexibility and Innovation | Less flexible to changes, innovation may be slower. | More adaptable to changes, encourages local innovation. | Decentralized model provides autonomy for rapid innovation at the node level. |
| Patient Privacy and Confidentiality | Potentially more robust privacy protection mechanisms. | Privacy depends on individual node's adherence to standards. | Centralized models often have stronger, uniform privacy controls. |
| System Resilience and Reliability | Vulnerable to system-wide outages or failures. | Distributed nature offers higher resilience to systemic failures. | Decentralized systems are generally more robust against widespread system failures. |
| Interoperability and Integration | Easier internal system integration, challenges in external interoperability. | Requires standardized protocols for seamless interoperability. | Centralized models excel in internal data integration but may struggle with external data sources. |
| Data Integrity and Quality Control | Easier to enforce data standards and maintain data integrity. | Data integrity reliant on individual node's policies and practices. | Centralized models can more efficiently monitor and audit data integrity. |
| Update and Maintenance Efficiency | Uniform system updates, centralized maintenance. | Updates and maintenance must be managed individually across nodes. | Centralized systems allow for streamlined updates and maintenance processes. |

interaction of the framework with actual healthcare data and systems, highlighting its adaptability and efficiency in managing privacy and security concerns.

- Empirical Research and Comparative Analysis:** A significant aspect of the evaluation process is conducting empirical research and comparative analyses. By benchmarking CEMPS against existing EHR models, as suggested in studies such as [18] and [19], the relative strengths of the framework and the areas for improvement are identified. This comparative approach not only reinforces the position of the framework in the current EHR landscape, but also provides valuable information for future improvements.
- Performance Metrics and KPIs:** The effectiveness of the framework is also measured through various performance metrics and Key Performance Indicators (KPIs), resonating with the approaches described in [242] and [243]. Metrics such

as data retrieval speed, data processing accuracy, and the incidence of security breaches are meticulously tracked and analyzed. These quantitative measures provide objective data on the operational effectiveness of CEMPS, ensuring its alignment with the demanding standards of healthcare data management.

The CEMPS framework, as depicted in **Figure 3.11**, demonstrates a structured and layered approach, with the Evaluation Layer being integral to its overall integrity and effectiveness. Furthermore, the stages involved in the definition of the privacy model within CEMPS, as shown in **Figure 3.10**, emphasize the systematic and methodical nature of the development of the framework.

The Evaluation Layer ensures that the CEMPS framework is not only theoretically robust but also practically effective in addressing the dynamic challenges of modern healthcare data ecosystems. Through a combination of statistical analysis, empirical research, and performance metrics, this layer validates the capacity of the framework to revolutionize EHR systems, ensuring increased security, enhanced privacy, and improved operational efficiency.

3.6.1 Comparison with Existing Privacy Preservation Techniques

Despite the significant advances introduced by the CEMPS framework in enhancing the privacy and security of electronic health records, it is essential to contextualize its innovations within the broader spectrum of existing privacy preservation techniques in healthcare. This comparison aims to elucidate the distinctive features and advantages of CEMPS in relation to established methods.

3.6.1.1 Traditional Privacy Preservation Techniques

Traditional techniques for preserving privacy in healthcare data often rely on methods such as data anonymization, pseudonymization, and encryption. While these methods provide foundational privacy safeguards, they frequently encounter limitations in scenarios involving complex data integration and real-time data access requirements.

3.6.1.2 Advanced Privacy-Enhancing Technologies (PETs)

Recent developments in Privacy-Enhancing Technologies (PETs), including differential privacy, secure multiparty computation, and blockchain, offer more robust privacy guarantees. These technologies enable more secure data sharing and analysis without compromising individual privacy. However, the implementation of these PETs in healthcare

settings can be challenged by scalability issues and the complexity of healthcare data ecosystems.

3.6.1.3 CEMPS Framework Advantages

The CEMPS framework distinguishes itself by integrating the principles of Federated Learning (FL) and Differential Privacy (DP) to address both the security concerns and the practical usability of EHR systems. Unlike traditional methods, which often operate in isolation, CEMPS provides a comprehensive approach that enhances data privacy and security while maintaining data utility and interoperability. Additionally, the framework's focus on scalable architecture and compliance with healthcare regulations further underscores its suitability for modern healthcare ecosystems.

By drawing on the strengths of FL and DP, CEMPS effectively mitigates the risks associated with data re-identification and unauthorized access, offering a more adaptable and resilient solution for healthcare data privacy and security. This comparison highlights CEMPS's contribution to advancing privacy preservation techniques in healthcare, offering a path forward for securely managing sensitive health data in an increasingly digital world.

3.6.2 Critical Evaluation of CEMPS

Following this comparison, it is crucial to critically evaluate the CEMPS framework's performance against these existing techniques. This evaluation will consider various metrics, including privacy preservation efficacy, scalability, adaptability to healthcare regulatory changes, and ease of integration with existing healthcare IT ecosystems.

The table 3.2 compares traditional and advanced privacy preservation techniques in healthcare, highlighting their foundational principles, scalability, regulatory compliance, and challenges. Techniques range from anonymization and encryption to more sophisticated approaches like Differential Privacy and Blockchain. It places special emphasis on the CEMPS framework, which integrates Federated Learning and Differential Privacy, showcasing its superior scalability, robust compliance with privacy laws like GDPR and HIPAA, and effectiveness in tackling re-identification risks and unauthorized access, positioning CEMPS as an innovative solution for healthcare data privacy and security.

TABLE 3.2: Comparative Analysis of Privacy Preservation Techniques in Healthcare

| Privacy Technique | Basis of Technique | Applicability in Healthcare | Scalability | Regulatory Compliance and Challenges |
|---|---|--|--|---|
| Traditional Methods (Anonymization, Pseudonymization, Encryption) | Data masking and encryption | Suitable for basic EHR systems | Limited by static data structures | Compliance with basic privacy regulations; Risk of re-identification |
| Advanced PETs (Differential Privacy, Secure Multiparty Computation, Blockchain) | Mathematical guarantees, decentralized ledger | Emerging applications in EHR and health data exchanges | Varies by technology; some are more scalable than others | Stronger privacy guarantees; Complexity and integration challenges |
| CEMPS Framework | Federated Learning and Differential Privacy | Comprehensive EHR systems with focus on privacy and security | High scalability and adaptability to large datasets | Enhanced compliance with GDPR, HIPAA; Addresses re-identification and unauthorized access risks |

3.7 Limitations

The CEMPS framework, while showcasing significant strides in improving privacy and security within Electronic Health Records (EHR) systems, encounters several challenges that are critical to address for its broader applicability and effectiveness. This section defines these challenges and underscores the need for ongoing refinement and development of the framework.

1. **Scalability Concerns:** As healthcare data volumes continue to grow at an exponential rate, the scalability of the CEMPS framework becomes an essential concern. Ensuring effective scalability while maintaining a delicate balance between robust privacy/security measures and optimal system performance is a complex endeavor. This scalability is essential to meet the diverse and evolving needs of healthcare providers and patients, as emphasized in studies such as [260] and [239].
2. **Standardization of Data Formats:** The heterogeneity of data formats in various healthcare entities poses a significant challenge to the uniform application and effectiveness of the CEMPS framework. Obtaining standardization, despite the inherent diversity of healthcare settings, is crucial for the success of the framework, but it remains a challenging task, highlighted in works such as [240] and [241].
3. **Risk of Data Re-identification:** Despite the incorporation of advanced techniques such as differential privacy (DP), CEMPS cannot entirely eliminate the risk of data re-identification. This risk is particularly acute in scenarios dealing with high-dimensional data or potential cross-referencing across multiple datasets. The need to continuously improve privacy techniques within the framework to mitigate these risks is discussed in [18] and [19].

The dynamic landscape of technology and healthcare requires that frameworks like CEMPS remain agile, adaptable, and responsive to emerging challenges. The integration of novel technologies and methodologies is critical in this regard. Additionally, the insights of this research contribute significantly to the ongoing discourse on privacy and security in EHR systems. They call for a proactive approach to anticipating and addressing potential challenges and limitations.

In summery, while CEMPS represents a significant advancement in the realm of EHR system privacy and security, its continuous evolution is essential. Addressing its current limitations and adapting to the rapidly changing healthcare and technological landscape will ensure its sustained efficacy as a leading solution for secure and private healthcare data management.

3.8 Future Work

The development path of the CEMPS framework is geared towards addressing its current limitations and adapting to the evolving landscape of healthcare data management. This section outlines the primary areas of focus for future work, emphasizing the enhancement of the framework's scalability, data standardization, and security measures.

1. ***Scalability and Data Standardization:*** Efforts will be directed towards amplifying the scalability and data standardization capabilities of CEMPS. This enhancement is vital for managing large datasets and ensuring seamless interoperability across various healthcare systems, as highlighted in studies such as [249] and [250].
2. ***Streamlining Implementation:*** A key focus will be to streamline the implementation process, reducing the dependency on technical expertise and improving the accessibility and usability of the framework. This approach aims to make CEMPS more universally applicable, as suggested in [251] and [261].
3. ***Updating Compliance Protocols:*** Continuous updates of compliance protocols in line with evolving healthcare legislation will be a priority. Additionally, strategies will be enhanced to prevent data re-identification and safeguard patient privacy, a concern highlighted in [111] and [258].
4. ***Advanced Security Measures:*** Future development will also involve exploring and integrating advanced security measures to protect against emerging cyber threats. This aspect is crucial to maintaining the integrity and confidentiality of EHR systems, as discussed in [259] and [260].

Future research directions informed by these findings should focus on exploring innovative strategies to further improve scalability, data standardization, and privacy protection techniques within the CEMPS framework. These efforts are expected to not only reinforce the framework, but also contribute significantly to the development of more resilient and efficient EHR systems worldwide, as envisioned in current research.

In essence, the CEMPS framework stands as a critical advance in the domain of EHR systems, marking a significant leap in the management of sensitive health data. Its future development is poised to realize a more interconnected, efficient, and secure EHR system, embodying a comprehensive approach to data management that is firmly anchored in privacy and security considerations.

3.9 Conclusion

This chapter has meticulously defined the development and implementation of the Centralized EHR Model for Preserving Privacy and Security (CEMPS), a transformative framework poised to revolutionize EHR data sharing within the healthcare ecosystem. At its core, CEMPS is engineered to facilitate a secure and privacy-compliant exchange of sensitive personal health information among diverse healthcare stakeholders, representing a paradigm shift in how healthcare data are managed and protected.

The development of CEMPS was underpinned by a comprehensive exploration of existing research challenges in the realm of EHR systems. Through an extensive literature review, this research identified critical gaps in current methodologies, particularly in the areas of data privacy, security, and interoperability. This thorough investigation was instrumental in shaping the strategic development of CEMPS, ensuring that the framework not only addresses current needs but also anticipates future demands in healthcare data management.

A key accomplishment of CEMPS is the establishment of a robust and resilient framework that guarantees the security and privacy of health data. This framework, crafted through rigorous development, implementation, and continuous evaluation processes, conforms to the highest standards of data protection and cyber resilience. The integration of advanced technologies and methodologies within CEMPS underscores its commitment to protecting patient information. Furthermore, the framework has been designed with the agility to adapt to the ever-evolving landscape of healthcare technology, ensuring its relevance and efficacy in the face of new challenges and opportunities.

Furthermore, CEMPS contributes to the broader discourse on health data management by providing a blueprint for balancing security and accessibility. Its design philosophy

encapsulates a subtle understanding of the multifaceted nature of EHR systems, acknowledging the diverse needs and perspectives of various stakeholders, including healthcare providers, patients, and regulatory bodies. By offering a model that harmonizes these diverse interests, CEMPS sets a precedent for future innovations in the field.

The implementation of CEMPS also offers valuable insight into the practical aspects of deploying such a comprehensive system in a real-world setting. The lessons learned from this implementation can guide the development of similar frameworks in other sectors where data privacy and security are paramount.

Ultimately, CEMPS stands as a testament to the potential of innovative, technology-driven solutions to address complex challenges in healthcare data management. It exemplifies a forward-thinking approach to EHR systems, prioritizing the protection of patient data while improving the efficiency and effectiveness of healthcare services. The development journey of the framework, marked by meticulous research and strategic planning, paves the way for a new era in healthcare data sharing, a landscape characterized by enhanced security, enhanced privacy, and optimized operational efficiency.

Chapter 4

INTEGRATING ADVANCED TECHNOLOGIES AND ONTOLOGY MODELS FOR ENHANCED SECURITY IN ELECTRONIC HEALTH RECORDS

NOTE: The content of [Chapter 4, Part A](#) has been submitted to *IEEE Open Access Journal*.

Nowrozy, R., Ahmed, K., Wang, H., (2023, July). GPT, Ontology, and CAABAC: A Tripartite Personalized Access Control Model Anchored by Compliance, Context and Attribute, *IEEE Open Access Journal*. **Currently undergoing the second round of minor revisions, pending resubmission.**

NOTE: The content of [Chapter 4, Part B](#) has been accepted and published by *Health Information System Conference, Melbourne*

Nowrozy, R., Ahmed, K., Wang, H., (2023, July). Enhancing Health Information Systems Security: An Ontology Model Approach, *Health Information System Conference, Melbourne, 2023*. **Accepted and published.**

Introduction

This chapter offers an in-depth exploration of [central theme], divided into two complementary parts. [Part A](#), "*GPT, Ontology, and CAABAC: A Tripartite Personalized Access Control Model*", focuses on developing a novel access control model for Electronic Health Record (EHR) security. Introduces the GPT-Onto-CAABAC framework, integrating Generative Pretrained Transformer (GPT) technology, medical-legal ontologies, and Context-Aware Attribute-Based Access Control (CAABAC). This model is designed for personalized access control, combining proactive decision-making and compliance auditing with specific legal and healthcare considerations.

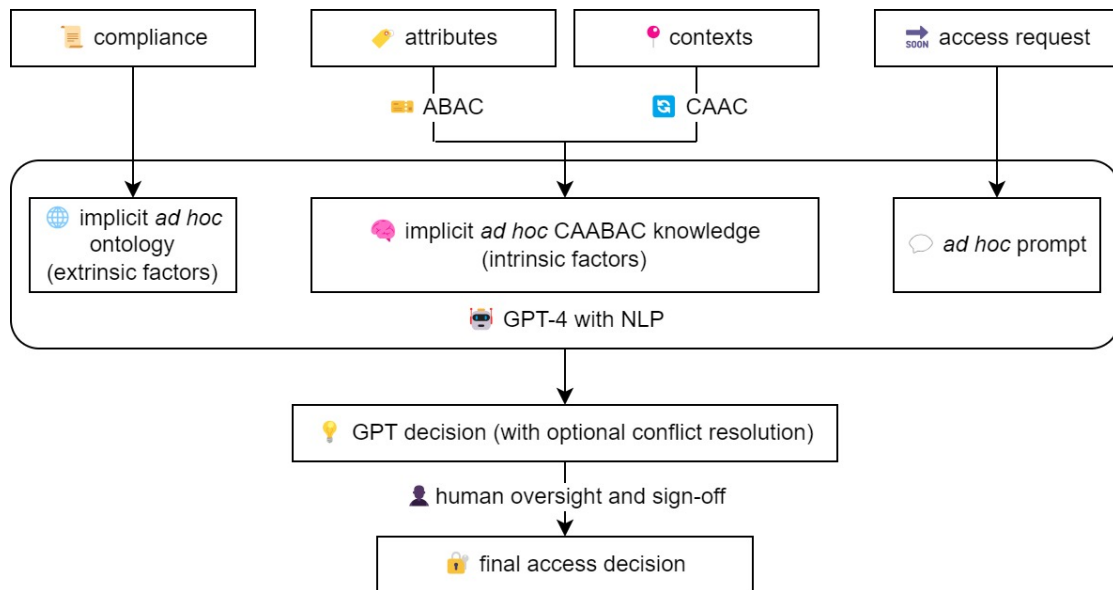
[Part B](#), "*Enhancing Health Information Systems Security: An Ontology Model Approach*," shifts attention to the broader issues of EHR security and privacy. It critiques current research limitations and proposes a holistic security ontology model. This model, grounded in the Conceptual Ontology Security Model, blends various security dimensions, namely confidentiality, integrity, availability, and access control strategies such as RBAC, ABAC, and MAC. The model aims at flexible and robust protection of health information, adhering to relevant regulations and policies.

The structure of the chapter, which progresses from [Part A](#) to [Part B](#), offers a layered examination of the security of EHRs, from specific access control models to broader security frameworks. This sequential approach ensures a comprehensive understanding of both individual innovations in EHR access control and the overarching challenges in the security of health information systems.

PART A: GPT, ONTOLOGY, AND CAABAC: A TRIPARTITE Personalized ACCESS CONTROL MODEL ANCHORED BY COMPLIANCE, CONTEXT AND ATTRIBUTE

NOTE: The content of this chapter has been submitted to PLOS ONE, and is currently under Peer Review. Nowrozy, R., et al. (2023, July). GPT, Ontology, and CAABAC: A Tripartite Approach to EHR Access Control Decisions, *PLOS ONE Journal* (<https://journals.plos.org/plosone/>).

Graphical Abstract



4.1 Introduction

The advent of *Electronic Health Records* (EHRs) has revolutionized healthcare by digitizing traditional paper-based records and centralizing patient data [288, 289]. These

digital systems have simplified administrative tasks [290, 291], improved clinical decision making [292, 293], and reduced medical errors [294, 295]. The integration of predictive analytics powered by *Artificial Intelligence* (AI) and machine learning has further enhanced treatment plans and patient outcomes [23]. During the COVID-19 pandemic, EHRs played a crucial role in monitoring viral spread, tracking patient outcomes, and accelerating research [56, 288, 296].

Despite these advancements, EHR systems face significant challenges in ensuring access control to maintain privacy and confidentiality. This balancing act is critical in enabling accessibility for healthcare professionals while complying with legal and ethical guidelines. Data breaches or misuse can have serious repercussions [53, 297]. Healthcare information, being a prime target for cyber threats, requires robust security measures [53, 121].

In response to these challenges, this chapter introduces the GPT-Onto-CAABAC framework, a novel tripartite access control model designed to enhance EHR security. This framework combines the strengths of *Generative Pretrained Transformer* (GPT), complex medical-legal ontologies, and the precision of *Context-Aware Attribute-Based Access Control* (CAABAC). It provides personalized access control decision recommendations by aligning with legal adherence, healthcare attributes, and patient environments. The GPT-Onto-CAABAC framework is central in offering personalized access control advice, facilitating both proactive decision-making and rigorous post-decision audits to ensure compliance with regulations. This innovative approach integrates the precision of ontology and access control systems with the flexibility of GPT for in-depth policy interpretation, showcasing significant potential in EHR access control flexibility and adaptability. Our evaluation indicates that this framework not only excels in the healthcare sector but also has broad applicability across various industries requiring access control decisions aligned with compliance and environmental considerations [6, 53].

Security breaches in 2022, which led to the exposure of sensitive data of over 20 million individuals in the USA, highlight the urgency of enhancing EHR security ¹. Fig. 4.1 shows the increasing trend of large data breaches in EHRs across the USA from 2008 to 2022 ². The successful use of ChatGPT-4 in business consulting, as indicated by the Harvard Business School study, emphasizes the potential for AI in sectors like EHR access control auditing ³. However, the industry's response to these security challenges has been lacking [6, 53], underlining the need for innovative solutions like our GPT-Onto-CAABAC framework.

¹<https://www.chiefhealthcareexecutive.com/view/the-11-biggest-health-data-breaches-in-2022>

²<https://www.healthit.gov/data/quickstats/office-based-physician-electronic-health-record-adoption>

³<https://www.afr.com/work-and-careers/workplace/consultants-using-ai-do-better-especially-underperformers-study-20230922-p5e6vi>

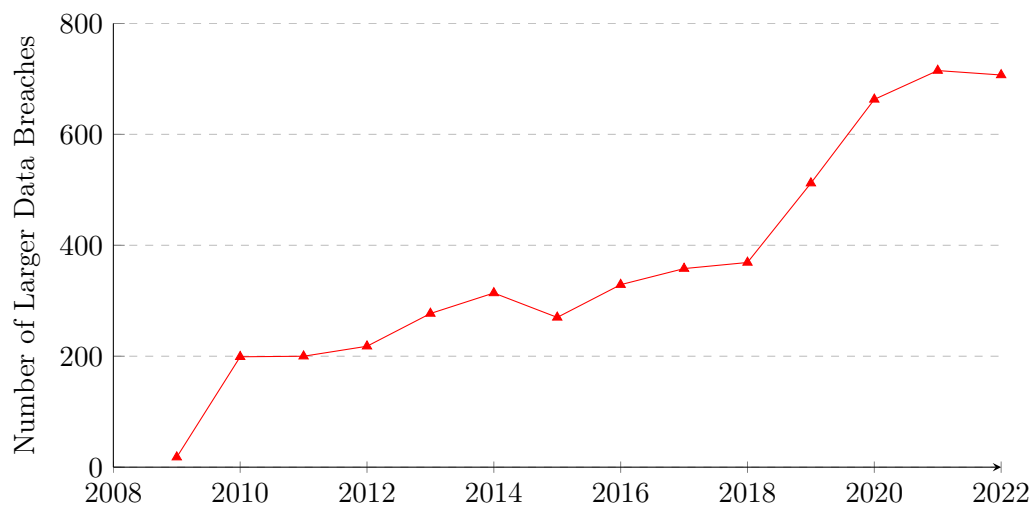


FIGURE 4.1: Number of Larger Data Breaches (≥ 500 Records Per Breach) of EHR from 2009 to 2022 in USA

Current models for EHR access control such as *Role-Based Access Control* (RBAC), *Attribute-Based Access Control* (ABAC), and *Context-Aware Access Control* (CAAC), although useful, present distinct challenges in adapting to dynamic healthcare settings [298–300]. The inflexibility of RBAC’s role-centric structure curtails its versatility, whereas ABAC and CAAC, while more adaptable, face operational challenges due to the complexity of managing attributes and the difficulty in defining and capturing context, respectively. Furthermore, current solutions aimed at addressing EHR interoperability issues, such as ontology-based methods, are not without difficulties. These methods struggle with the issues of data harmonization and semantic heterogeneity and often fail to consider organizational and cultural barriers to interoperability [270, 301, 302]. Despite considerable attempts to streamline and enhance these models, their inherent limitations in coping with the dynamic complexity of healthcare environments remain a concern. These limitations underscore the need for an innovative approach to EHR security that can integrate the strengths and address the shortcomings of existing models.

The transformative *Natural Language Processing* (NLP) capabilities of *Generative Pre-trained Transformers* (GPTs) have opened new horizons for the access control decision-making process [303]. Using GPT’s proficiency for real-time personalized recommendations and its subtle interpretation of multifaceted legal and ethical standards, we introduced the *GPT-powered Ontology-Driven Decision of Context-Aware Attribute-Based Access Control* (GPT-Onto-CAABAC) model [304–307]. This model embodies the collective strengths of *Context-Aware Attribute-Based Access Control* (CAABAC) and ontology-driven decision-making. The resulting framework is both adaptive and detailed. Central to this process is the establishment of context, devising an ontology congruent with

healthcare care norms, associating the context with the said ontology, formulating access policies, employing CAABAC, and finally rolling out the ontology-driven decision system. This holistic strategy fortifies data security. Our GPT-Onto-CAABAC model outperforms conventional retrieval-based systems by proficiently maneuvering through ever-shifting EHR access control scenarios. Addresses the rigidity of laws while accommodating the dynamism intrinsic to routine healthcare settings. Although our model exhibits strong potential to fortify EHR security, mitigate risks associated with data breaches, and acclimate to the evolving environment of healthcare settings, it also has broader implications. Although our focus remains tied to EHR access control scenarios, given their complicated compliance, malleability, and auditing stipulations, the approach has vast potential for access control decision auditing in varied contexts. The synergy of advanced NLP capabilities with structured access control models promotes an in-depth analysis that transcends healthcare, extending to any access control environment characterized by layered regulations and policies. The integration of GPT's NLP strengths with time-tested techniques such as ontology, CAAC, and ABAC facilitates the creation of complex policy-to-legal-ontologies. Moreover, it stimulates comprehensive collation of contextual details via CAAC and attribute information through ABAC, ensuring balanced access control decisions that heed the complexities of medical situations and EHR decision-making paradigms. Currently in its nascent, proof-of-concept stage, our GPT-Onto-CAABAC model holds promise as a transformative agent in both healthcare and diverse sectors, paving the path for a more cyber-resilient future.

The major contributions of our study include:

- 1) *Problem Analysis* (Section 4.3): a detailed analysis of the challenges and intricacies involved in access control decisions for electronic health records (EHRs), to highlight the limitations of existing systems and underscores the need for a more robust and context-aware solution.
- 2) *Innovative Solution* (Section 4.4): the proposed GPT-Onto-CAABAC framework, which combines GPT, ontology, and access control models for enhanced access control management in healthcare settings, with details on the high-level architecture and underlying components of the framework.
- 3) *Comprehensive Evaluation* (Section 4.5, 4.6): an exhaustive empirical analysis of our GPT-Onto-CAABAC framework in various healthcare contexts, using targeted metrics to assess real-world applicability, performance, and insights gained.

The rest of the chapter is organized as follows. Section 4.2 provides an in-depth review of related work in the field of access control systems. Section 4.3 introduces our theoretical framework GPT-Onto-CAABAC, which unites ontology, CAABAC, and the role

of GPT. Section 4.4 discusses our experimental design. Section 4.5 presents the findings and insights of our experiment. Section 4.6 looks at an insightful discussion of our results, including its limitations. Finally, Section 4.7 summarizes the research and outlines potential future directions.

4.2 Related Works

In the related work section, we review how access control models and ontology have been applied to make EHR access control decisions and their inadequacies.

4.2.1 Access Control in EHR

Access control is a fundamental aspect of security in information systems. In recent years, a myriad of studies have been conducted focusing on RBAC, ABAC, CAAC, and Ontology-based interoperability to address the various security concerns prevalent in EHRs [308]. However, these models often struggle to adapt to the complex, real-time decision-making required in healthcare settings, despite their inherent strengths.

4.2.1.1 RBAC in EHR Security

RBAC assigns permissions based on predefined user roles, offering a structured approach to EHR security that has garnered substantial academic interest [298, 309]. However, this model often falls short in dynamic healthcare settings. In particular, many studies [45, 298, 309–315] failed to adequately address the complexity of access control to the EHR, exhibiting deficiencies such as the lack of robust auditing mechanisms, insufficient granularity of user roles and permissions, and failure to adapt to emerging vulnerabilities and security threats. Additionally, aspects of RBAC such as role hierarchies, scalability, and implications of cloud-based EHR data storage have frequently been overlooked [314, 315]. These observations indicate the need for a more comprehensive strategy to address the practical utility and efficacy of RBAC in the security of EHR access control.

4.2.1.2 ABAC in EHR

The transition to ABAC models provided an additional layer of granularity and improved flexibility in EHR security [139]. However, the management of numerous attributes in large healthcare institutions with continuously evolving attributes posed challenges

[299, 316]. Significant deficiencies were also observed in the studies [15, 212, 297, 317–322]. These limitations mainly involved incomplete discussions on scalability, security vulnerabilities, practical considerations for EHR systems, efficient attribute management, and integration into existing healthcare systems. Therefore, more research is required to ensure a robust and effective implementation of ABAC in EHR security.

4.2.1.3 CAAC in EHR

The CAAC model enhanced the dynamic approach by incorporating contextual information [323]. However, capturing contextual information accurately and promptly posed a significant challenge due to the rapidly changing healthcare environment [324, 325]. Several implementations of CAAC exhibited weaknesses, especially in the area of EHR access control security [326–330]. Common limitations included a lack of comprehensive evaluations, a failure to address potential privacy and security concerns, insufficient detail on technical implementations, and a lack of real-world deployment evaluations. Therefore, while CAAC models show promise, more research is essential to address these challenges in their application to the security of EHR access control.

4.2.2 Ontology in EHR Security

The potential of ontology in access control of the EHR has been extensively investigated, yet revealed several limitations. [331] and [332] exposed the challenge of creating and maintaining comprehensive ontologies due to evolving healthcare standards, the lack of standardization, and the complex nature of healthcare data, which hampered interoperability and data sharing. Scalability issues and the complexity of managing complex access control policies were highlighted by [333] and [334]. These challenges increased when managing complex relationships, contextual information, and efficient searches for encrypted data in large-scale healthcare systems. [335] and [336] questioned the ability of ontology-based access control to capture the dynamic and context-dependent nature, handle granularity, or adapt to evolving user roles and temporal constraints. [337] emphasized the difficulty in maintaining comprehensive ontologies for the Circle of Care (COC) due to ever-changing healthcare settings. [125] developed an ontology and machine learning-based approach to enhance privacy in EHRs, aiming to balance privacy and accessibility while considering legal compliance, user-friendliness and cultural and social aspects, but their research was limited by the lack of comprehensive evaluation of the proposed model, including comparative analysis with other state-of-the-art approaches, scalability, and performance testing. Despite the potential of ontology-based

approaches in EHR access control, its application has encountered different but significant limitations, necessitating further research for its effective implementation.

4.2.3 Summary

Traditional access control models, despite their applicability in the healthcare sector, such as RBAC, ABAC, CAAC, and ontology-based access control [338–340], have proven essential for EHR security. However, they have faced significant challenges (Table 4.1). RBAC’s main hurdles include its rigidity in evolving healthcare settings, its limited granularity, and scalability problems [45, 298, 309, 315]. Although ABAC offers superior control, it creates complexity and requires resource-heavy operations in expansive, dynamic systems [139, 299, 316]. Comprehensive assessments and integration challenges are equally pressing [15, 212, 297, 317, 318]. CAAC’s ability to incorporate context into access requests is especially beneficial for the dynamic nature of healthcare care [300]. However, gathering precise, up-to-date context information becomes challenging due to rapid environmental changes [324, 325, 341]. Evaluation, applicability, and concerns about privacy further restrict its use [326–328]. The ontology-based access control model has encountered notable barriers, especially to maintain extensive ontologies with changing healthcare standards and to handle complex healthcare data [331–337].

Those traditional models have not wholly satisfied the security needs of access control in complex and dynamic environments, particularly in healthcare [342–345]. By contrast, our proposed GPT-Onto-CAABAC framework seeks to redress these deficiencies and has significant potential to bolster access control auditing across diverse industries. Thus, the need of the hour is research that ventures beyond healthcare, examining the framework’s utility in various highly regulated and dynamic scenarios. Future research efforts should amalgamate the adaptability of CAAC, the flexibility of ABAC, and the structure of RBAC while confronting novel threats, refining granularity, enhancing comprehensive auditing, fortifying authentication, refining attribute management, and ensuring scalability. The overarching aspiration remains to craft a robust, thorough, and pragmatic access control system not only for healthcare but also for other complex sectors.

4.3 Proposed Framework: GPT-Onto-CAABAC

In this section, we introduce our proposed framework: GPT-Onto-CAABAC (Fig. 4.2). Medical access control decision making balances both inflexible legal parameters and flexible daily situations that demand adaptability and context awareness. Given this

TABLE 4.1: Comparison of different access control models in addressing extrinsic and intrinsic factors (✓: capable; △: partially capable; ×: incapable)

| Access control models | Extrinsic factors | Intrinsic factors | |
|-------------------------------------|-------------------|-----------------------|----------------|
| | | Environmental context | Access subject |
| Traditional access control | RBAC | ✓ | × |
| | ABAC | △ | × |
| | CAAC | × | ✓ |
| Ontology | | △ | △ |
| <i>GPT-Onto-CAABAC (This study)</i> | | ✓ | ✓ |

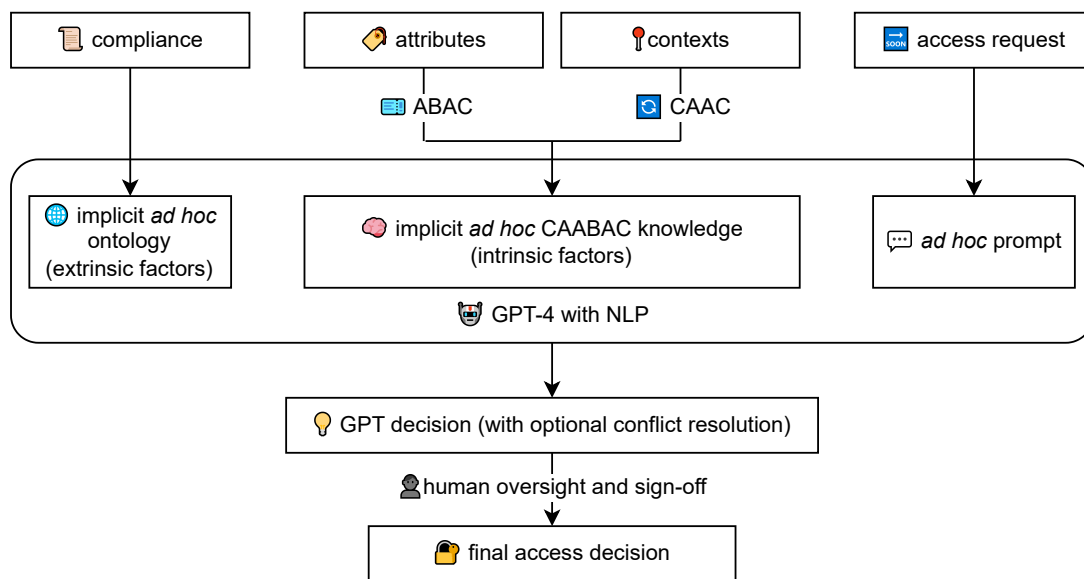


FIGURE 4.2: GPT-Onto-CAABAC

elaborate blend of static and dynamic elements, this study looks into the critical convergence of Ontology, CAABAC, and the transformative influence of GPT.

4.3.1 High-level framework overview

Our GPT-Onto-CAABAC framework serves as an integrated and versatile model for auditing access control decisions in various contexts. In particular, it adeptly addresses healthcare’s complex blend of compliance, flexibility, and auditing needs. By amalgamating ontology, CAABAC, and GPT, this framework demonstrates its unique prowess in dynamic and context-aware EHR access control. The framework’s components, as such, position it as exceptionally well-suited for post-decision audits in complex settings governed by multifaceted regulations. Initiating its process, the framework harnesses GPT’s capabilities to internally construct an implicit, transient ontology from legal texts and policies. This implicit *ad hoc* ontology model, unlike traditional ontologies, remains embedded within the GPT layer during runtime. This approach bypasses resource-intensive

Algorithm 1 GPT-Onto-CAABAC Process with Human Oversight

Require: Legal texts and policies \mathcal{P}

Require: Context information \mathcal{C}

Require: GPT model \mathcal{G}

- 1: $\mathcal{O} \leftarrow f_{\text{extraction}}(\mathcal{P})$ {Transform established policies to ontology}
 - 2: $\mathcal{A} \leftarrow f_{\text{capture}}(\mathcal{C})$ {Capture and standardize context with CAABAC}
 - 3: $D \leftarrow f_{\text{decision}}(\mathcal{O}, \mathcal{A}, \mathcal{G})$ {Initial decision making with GPT}
 - 4: **if** conflicts in D **then**
 - 5: $D' \leftarrow f_{\text{resolution}}(D, \mathcal{O}, \mathcal{A}, \mathcal{G})$ {Resolve conflicts with GPT}
 - 6: **else**
 - 7: $D' \leftarrow D$ {No conflicts, keep initial decision}
 - 8: **end if**
 - 9: $D_f \leftarrow f_{\text{human}}(D')$ {Human oversight and final sign-off}
 - 10: **return** D_f {Final decision}
-

ontology management, but lays a solid foundation for rule formulation and compliance [346, 347]. Subsequent to this implicit ontology formation, the model captures the real-time context and maps it to an *ad hoc* CAABAC model. By incorporating the attributes of users, resources, and the environment, it refines access decisions and customizes them to distinct needs [348]. The GPT layer within the framework is tasked with dynamic decision making. It reconciles potential conflicts between context- and policy-based rules while ensuring strict conformity to legal and institutional frameworks, thus improving system accountability and credibility [307].

Our multicomponent approach is represented by Algorithm 1, which details the interaction of each element to yield informed and compliant access control decisions. By transcending the limitations of existing models, this innovative framework adjusts access control based on various situational factors and remains rooted in regulatory mandates [348, 349]. The fusion of ontology precision, CAABAC adaptability, and GPT generative prowess gives birth to the GPT-Onto-CAABAC model, portraying a flexible but methodically structured access control mechanism [307]. This framework is poised to guide the evolution of healthcare data security approaches, proposing a solution that is robust and adaptable to subtle contextual factors.

4.3.2 Detailed ontology explanation

Ontology in access control serves as a structured knowledge representation, cataloging distinct entities and defining their associated properties and interrelationships [309, 332]. This structured approach is vital for the conversion of high-level policies into executable rules, which form an indispensable element of the decision-making apparatus in complex operational settings [332]. Simultaneously, CAABAC employs a detailed approach to access control that takes into account various user attributes within specific contexts.

This allows for the generation of precise and adaptable access control decisions [323]. By addressing the limitations and leveraging the strengths of both, our framework pioneers an innovative ontology. This new ontology represents a complex network of relationships between various contextual elements and user attributes while also providing a clear framework for decision-making processes. It also integrates seamlessly with the CAABAC mechanisms, creating an enriched access control model [323].

In healthcare settings, ontologies function as explicit formal specifications for domain-specific entities and their interconnections [350, 351]. They offer a consistent and structured interpretation of inflexible access control components such as laws, regulations, and policies. The notion of a medical-legal ontology encapsulates these fixed components, facilitating efficient data retrieval, management, and query execution while ensuring that the system remains compliant with legal requirements [350]. The efficacy of access control models in EHRs is influenced by both external factors, such as laws, regulations, and institutional guidelines [352, 353], and internal factors that arise from the dynamic healthcare delivery environment [353]. Although existing models such as RBAC, ABAC and CAAC each have limitations in managing these complexities [353], our ontology-centered approach provides a balanced mechanism to manage these factors effectively. Compliance with external policies is ensured to comply with legalities and safeguard patient data, while adaptability to internal factors is addressed to improve system usability and operational efficiency.

The crucial transition of policies into a formal ontology employs NLP techniques to metamorphose unstructured legal verbiage into ontologies that are implicitly understood and ad hoc in nature to human experts, while remaining structured and machine-comprehensible for automated processing by GPT. This includes the identification of relevant entities, the mapping of relationships, and semantic parsing [350]. The resulting medical-legal ontology serves as a distilled representation of principles derived from these legal texts, thus establishing the operational limits for the system. Furthermore, as laws and policies evolve, this NLP capability enables an efficient update of the 'medical-legal ontology,' eliminating the need for manual reengineering prevalent in conventional ontology methods[48–50, 61].

$$\mathcal{O} = f_{\text{extraction}}(\mathcal{P}) \quad (4.1)$$

Here, \mathcal{P} denotes the policies, and \mathcal{O} symbolizes the resultant ontology. The function $f_{\text{extraction}}$ encapsulates the ontology extraction process.

4.3.3 Detailed CAABAC explanation

The CAABAC model amalgamates the merits of CAAC and ABAC to deliver an adaptive, fine-grained access management mechanism, especially suitable for healthcare settings.

4.3.3.1 Advantages of ad hoc contextual information in healthcare

One of the most compelling aspects of CAABAC lies in its ability to dynamically construct contextual ad hoc information for immediate consideration in access control decisions. This characteristic is highly relevant in healthcare settings for multiple reasons:

- **Temporal Sensitivity:** Rapidly evolving healthcare settings can have significant repercussions if access is delayed. Therefore, real-time contextual information is crucial.
- **Resource Efficiency:** One-off ad hoc contextual data prevent system clutter, optimizing resources for more urgent needs.
- **Enhanced Security:** Eliminating contextual information ad hoc after decision making minimizes the risks related to unauthorized access and data leakage.
- **Precision in Decision-making:** Instant contextual construction allows for highly tailored access control decisions, essential when handling sensitive health records.
- **Compliance and Auditing:** Contextual real-time information promotes better compliance with legal and ethical data access and privacy requirements. Immediate data disposal is consistent with the principle of data minimization.

This approach provides a balanced solution, advantageous in the complex, fast-paced, and regulated healthcare sector.

4.3.3.2 Role of CAAC

CAAC primarily addresses the dynamic and situational subtleties in access control by tailoring decisions to the existing contextual environment. Within healthcare, practitioners are often faced with a spectrum of contextual states that include emergencies, different patient statuses, and diverse technological ecosystems. CAAC navigates these variations effectively, abiding by the rules and constraints defined by the ontological framework. Consequently, this facilitates an increase in workflow efficiency while preserving data integrity and confidentiality.

4.3.3.3 Contribution of ABAC

In contrast, ABAC augments CAAC by incorporating a multifaceted attribute-based decision-making process. This allows attributes tied to users, resources, and the operational environment to be considered in decision making. These attributes can be highly specific, ranging from clinical flags like *Not For Resuscitation* (NFR) to device categories such as hospital-approved devices or *Bring Your Own Device* (BYOD). Thus, ABAC introduces a level of specificity that accommodates complex and multifaceted healthcare scenarios.

4.3.3.4 Distinction between CAABAC and ABAC

While ABAC is primarily attribute-centric, CAABAC leverages contextual awareness to provide a more adaptive and responsive access control mechanism. Unlike traditional ABAC, CAABAC dynamically adapts to situational changes, offering a higher level of granularity in access decisions, making it particularly beneficial in the dynamic and fluctuating environment of healthcare care provision.

4.3.3.5 GPT-Onto-CAABAC context capture

To accommodate this dynamicism, the GPT-Onto-CAABAC framework features a specialized context capture module. This subsystem harvests data from the Electronic Health Record (EHR) and the prevailing situation, transmuting these unstructured inputs into a set of standardized attributes consistent with the CAABAC model. Standardization accounts for multiple variables, such as user roles, ongoing tasks, objects involved, and environmental conditions. Health professionals can also contribute context or attribute data in natural language, which is then processed and understood by GPT for seamless integration into the decision-making process[62–66].

$$\mathcal{A} = f_{\text{capture}}(\mathcal{C}) \quad (4.2)$$

In Equation 4.2, \mathcal{C} symbolizes the context information, \mathcal{A} symbolises the standardized attributes used in CAABAC, and f_{capture} is the function responsible for contextual capture and standardization.

4.3.3.6 Example of CAABAC

Consider an emergency room scenario where a patient is admitted with a critical condition. Contextual factors include the emergency state, the critical health status of the patient, and the role of the treating physician. A nurse logs into the system to access the patient's medical history. In this scenario, ABAC attributes could include the role of the nurse, credentials, and the level of data sensitivity of the medical records. CAAC contextual information could involve real-time factors such as the emergency state, the urgency level coded by the attending physician, and the time-sensitive nature of the required data access. Integrating these, the CAABAC model dynamically grants access because the situation is deemed emergency and the nurse's role is verified as authorized to access critical health information in these specific circumstances. By adhering to these specifications, CAABAC not only meets, but enhances, the prerequisites for secure, adaptable, and fine-grained access control, specifically within the healthcare sector.

4.3.4 GPT integration and conflict resolution

GPT models excel in NLP tasks and human-like text generation, showcasing immense potential for deployment in diverse sectors, including healthcare [354–356]. Our framework aims to harness these capabilities to enhance ontology-based decision making and CAABAC in medical access control systems. Importantly, the GPT-Onto-CAABAC framework utilizes GPT models specifically for compliance checks and not for real-time access control decisions. The reason for this distinction is twofold: first, GPT models, while adept at complex language tasks, may have response generation times that render them unsuitable for time-sensitive healthcare scenarios; second, traditional access control models are more appropriate for real-time decisions due to their optimized speed and established reliability.

Integration with GPT equips the system with tools to resolve conflicts between ontology, CAAC, and ABAC. This includes interpreting the medical-legal ontology and offering resolutions within legal limits, considering the context and attributes involved. The self-improving nature of GPT also means that the model refines its recommendations over time, thus fortifying the resilience of the GPT-Onto-CAABAC model. In GPT-Onto-CAABAC, conflict resolution is crucial, where the ontology, which encapsulates legal and institutional frameworks, has primacy over CAAC and ABAC. However, CAAC and ABAC may overwrite each other within the bounds of the ontology, depending on the context and attributes. A well-structured conflict resolution mechanism ensures this delicate balance between security and usability[67–70].

The decision-making module employs GPT’s capabilities to generate detailed recommendations. Trained in the developed ontology and the CAABAC attributes, GPT enables the system to understand the complex interplay between static rules and the dynamic context. As a response to the reviewer’s feedback, the system not only grants or denies access but also suggests a range of contextually appropriate and policy-compliant actions. Unlike conventional binary access controls, this flexibility allows provisional granting of access under specific conditions, thereby satisfying both regulatory requirements and clinical needs. The mathematical formulations of this decision-making process are as follows:

$$D = f_{\text{decision}}(\mathcal{O}, \mathcal{A}, \mathcal{G}) \quad (4.3)$$

In scenarios where decision-making might introduce conflicts or ambiguities, a conflict resolution function is invoked.

$$D' = f_{\text{resolution}}(D, \mathcal{O}, \mathcal{A}, \mathcal{G}) \quad (4.4)$$

4.3.5 Human oversight and sign-off

The inclusion of AI in healthcare increases human capabilities, optimizes operations, and increases productivity [91, 357]. However, the GPT-Onto-CAABAC model further incorporates human oversight and final sign-off to acknowledge the indispensable expertise and judgment that healthcare professionals contribute. This integration is instrumental in maintaining ethical standards and ensuring the delivery of responsible health services [55, 358]. Although GPT and AI models are highly capable, they are limited in capturing the ethical subtleties and multifaceted decision-making inherent in human expertise. Human oversight is an important protective layer against inaccuracies or shortcomings inherent in automated decision-making processes [359]. AI models, although advanced, are susceptible to errors and require an additional layer of scrutiny from humans to preclude detrimental consequences and ensure patient safety. Furthermore, the presence of human supervision in the system increases public trust in technology, as it serves as a reassurance that decisions are validated by accountable professionals [55, 360]. The importance of human oversight serves to mitigate the risk of blindly accepting AI-generated decisions, which may lack depth of ethical or professional considerations. If a human mistakenly override an accurate GPT recommendation, a secondary review mechanism could be enacted that involves expert consultation or peer review, thus adding another layer of verification [361].

The GPT-Onto-CAABAC framework introduces a function, f_{human} , applied after the AI-based decision-making process, to allow human validation of AI-generated recommendations. Mathematically, the final decision D_f can be articulated as follows:

$$D_f = f_{\text{human}}(D') = f_{\text{human}}(f_{\text{resolution}}(D, \mathcal{O}, \mathcal{A}, \mathcal{G})) \quad (4.5)$$

In this equation, D_f denotes the ultimate decision, D' represents the initial decision of GPT, and \mathcal{O} , \mathcal{A} , and \mathcal{G} signify the ontology, attributes, and GPT model, respectively. The function f_{human} encapsulates human oversight and final validation, highlighting the commitment to ethically responsible AI and balancing technological capabilities with human expertise [362].

4.4 Implementation of the GPT-Onto-CAABAC framework

The efficacy of the GPT-Onto-CAABAC framework was evaluated through a series of carefully designed experiments, the results of which provide valuable information on its performance and potential improvements. This section outlines the design of our experiments, describing the datasets used, and the scenarios created to assess the GPT-Onto-CAABAC framework's capabilities. We have used the following steps to build our prototype.

- 1) *Construction of policy-to-legal-ontology* (Subsection 4.4.1): Import the 3 pieces of legislation into our ChatGPT-4-based model to build the policy-to-legal-ontology.
- 2) *Employment of Datasets* (Subsection 4.4.2): Use both real case studies and constructed scenarios as data sets.
- 3) *Obtaining Decisions and Recommendations* (Subsection 4.4.3): Use our custom-constructed prompt 2 (to give the example once we have it) to feed the improved case study with information required by CAAC and ABAC, into our legal ontology, to seek access control decision and, if denied, recommendation to obtain access approval.
- 4) *Human Evaluation and Sign-off* (Subsection 4.4.4): Evaluate the results using our evaluation metrics.

TABLE 4.2: List of legislations governing EHR access

| Legislation | Jurisdiction level | Current Version |
|----------------------------|--------------------|-----------------|
| Privacy Act 1988 | Federal | 1 Sep 2021 |
| My Health Records Act 2012 | Federal | 1 Sep 2021 |
| Health Records Act 2001 | State of Victoria | 2 Sep 2022 |

4.4.1 Construction of policy-to-legal-ontology

The construction of the legalontology policy involves identifying key laws and regulations relevant to the context of access to the EHR access. For our use case, we have focused on the legal framework within the State of Victoria in Australia, identifying three key pieces of legislation, as detailed in Table 4.2.

- **Privacy Act 1988**⁴: A comprehensive privacy law detailing principles around personal data collection, use, and disclosure.
- **My Health Records Act 2012**⁵: Establishes the My Health Record system, a national EHR system.
- **Health Records Act 2001**⁶: Defines patients' rights for health records access and health care providers' responsibilities.

We incorporated the legislations into our model using the AskYourPDF⁷ plugin of ChatGPT-4, which facilitated the import of published PDF versions of the legislation. We did not create an explicit, clear ontology model, which often proves too rigid and fails to fully capture the complex reality of healthcare scenarios comprehensively. Instead, we leveraged ChatGPT-4's ability to understand and retain the implications of the legislation, effectively embedding an implicit legal medical ontology within the model's attention and knowledge layers. Although unconventional, this methodology leverages the inherent flexibility of the GPT architecture, harnessing the strengths of explicit and implicit knowledge representation. Our approach was demonstrated as a proof-of-concept implementation on ChatGPT-4, utilizing its robust hardware and computing capabilities. The resulting implicit legal medical ontology, validated under human oversight, forms the cornerstone of our GPT-Onto-CAABAC model and serves as the initial step towards our ultimate goal of creating a domain-specific *Large Language Model* (LLM) trained on this ontology.

⁴<https://www.legislation.gov.au/Details/C2014C00076>

⁵<https://www.legislation.gov.au/Details/C2021C00475>

⁶<https://www.legislation.vic.gov.au/in-force/acts/health-records-act-2001/047>

⁷<https://askyourpdf.com/upload>

4.4.2 Utilisation of datasets

Our strategic approach involved the construction of a comprehensive data set comprising more than 120 use case scenarios in 12 categories to improve the precision and reliability of the GPT responses. This methodology has been indispensable for multiple reasons:

- **Diverse Dataset:** Incorporating various EHR-related scenarios diversified the dataset, enriching the GPT learning experience. This diversity facilitated the model in generalizing and making accurate predictions in real-world applications.
- **Comprehensive Coverage:** By curating a minimum of 10 specific use case scenarios for each category, the data set provided an extensive representation of potential healthcare sector interactions, capturing its inherent complexities.
- **Cross-Referencing Legal Frameworks:** We cross-referenced the scenarios with the Australian Privacy Act 1988 and My Health Records Act 2012, enabling GPT to grasp the legal consequences of various situations, thus increasing its capacity for legally compliant recommendations.
- **Enhanced Accuracy:** Leveraging a large, diverse data set stimulated an improvement in the accuracy of the GPT's responses by exposing it to a wide range of situations and subtle contexts.
- **Improved Experimental Process:** Employing an expansive dataset enriched the experimental process, offering a vast source of data for training, testing, and validation, thus strengthening the GPT model.

In our experiment, we utilized a combination of two datasets that served distinct purposes. The first dataset included anonymized real-world EHR data, providing our system with realistic data points. The second dataset consisted of carefully constructed artificial scenarios that targeted specific capabilities of the GPT-Onto-CAABAC framework. These scenarios, which incorporated instances of high-frequency access requests, complex contextual conditions, abrupt legal or policy changes, and conflicting policies or extraordinary medical situations, offered an opportunity to evaluate the framework's robustness and adaptability. The construction of this comprehensive dataset, which included 120 use-case scenarios in 12 categories, was instrumental in addressing concerns about the provision of practical examples and empirical data. This data set played a critical role in improving the accuracy, reliability, and legal compliance of GPT responses. The diversity of the dataset not only facilitated the model in making accurate predictions and generalizing across various scenarios, but it also enhanced its versatility. Furthermore, the alignment of the scenarios with the Australian Privacy Act 1988 and

My Health Records Act 2012 guaranteed the model’s ability to provide legally compliant recommendations. The incorporation of real-world EHR data and the tailored artificial scenarios were critical in assessing the model’s adaptability and robustness under diverse conditions, yielding invaluable insights into its performance. Consequently, our methodology provided a wealth of empirical data and practical examples, highlighting the versatility, adaptability, and legal compliance of the GPT-Onto-CAABAC framework. In sum, the carefully constructed dataset and the testing scenarios facilitated a rigorous examination of the model’s performance, validating its potential for practical applications in healthcare access control.

4.4.3 Acquiring decisions and recommendations

The GPT-Onto-CAABAC framework employs ChatGPT-4’s advanced NLP capabilities to derive access control decisions and provide recommendations. These decisions and recommendations are contingent upon two primary elements: non-negotiable policy-to-legal-ontology and negotiable context and attribute information. Both elements influence the model’s understanding of EHR access control scenarios and guide its decision-making process. The nonnegotiable policy-to-legal-ontology, founded on existing legal regulations and healthcare policies, constitutes a rigid baseline for decision making. It is indispensable to ensure adherence to pre-established privacy and security requirements in EHR data management. In this proof-of-concept stage, several strategic decisions are adopted for both practicality and exploratory value. Firstly, ChatGPT-4 is used in its commercial form, negating the need for retraining or fine-tuning. This decision allows for an assessment of the model’s capabilities in a generic setting and offers future implementers the latitude to add domain-specific optimizations. Second, the framework does not retain CAABAC information, but rather acquires it ad hoc for each evaluation. Such a design aligns well with the inherently dynamic and complex environment of the healthcare sector, enabling adaptive access control decisions based on real-time situations rather than rigid processes. Lastly, we deliberately abstain from optimizing the model’s response time at this stage. This leaves room for prospective organizations to make performance-based adjustments tailored to their specific requirements when scaling from a proof-of-concept to a full-fledged implementation.

The negotiable context and attribute information give the system the flexibility to adapt and respond to the dynamic, multifaceted nature of the healthcare sector. The model processes an access request by receiving a prompt that describes the scenario in natural language. This prompt serves as the interface through which the context and attribute information is encoded and absorbed by ChatGPT-4. For example, a typical prompt might state:

Request for patient John Doe's EHR for a clinical study by Dr.

John Smith, who has a security clearance. Is access granted?

Outputs based on such prompts could be categorised as follows:

- Access granted: "Access granted. Ensure to maintain data confidentiality."
- Access denied: "Access denied. This is illegal."
- Recommendations: "Need to seek patient's informed consent. Seek permission from the ethics committee for special ethics approval."

The model cross-checks this information against the embedded policy-to-legal ontology. The decision is influenced not just by this ontology but also by the specific context and attributes presented, thus utilizing a form of deductive reasoning. In instances where access is denied, the model proposes recommendations for altering the context or attribute information to facilitate potential access approval. These could range from seeking permissions from higher authority to modifying the timing or environment of access. Thus, the GPT-Onto-CAABAC framework effectively balances regulatory adherence with the necessary flexibility in navigating the complex landscape of the healthcare sector.

4.4.4 Human evaluation and sign-off

The results are presented for human evaluation and approval. During our evaluation, there is no need to sign off other than human inspection and oversight to evaluate the effectiveness of GPT decisions and recommendations. For evaluation, we need to establish quantitative metrics. These could include:

4.4.4.1 Compliance

Measures the rate at which the system's decisions align with existing rules and policies. This could be calculated by identifying instances where the system's decisions were compliant with the rules and policies divided by the total number of decisions made. For example, if in 100 decisions, 95 were compliant with the policies, the compliance rate would be 95%.

4.4.4.2 Adaptability

Calculate how quickly the system adapts to sudden changes in policies or rules. This would ideally be measured over a period of time following the implementation of new rules or policies. You would compare the system performance (in terms of compliance rate, efficiency, and recommendation quality) immediately after the change and after a certain period, say, one month. The adaptability score could be the rate of improvement in system performance during this period.

4.4.4.3 Conflict Resolution Efficiency

Evaluates how effectively the system resolves conflicts between different policies or rules. This could be determined by identifying cases where there was a conflict between policies or rules and seeing how often the system made the correct decision. If there were 50 conflict cases and the system resolved 40 correctly, the efficiency of conflict resolution would be 80%.

4.4.4.4 Recommendation Quality

The evaluation of the quality of the recommendation requires a detailed analysis of the competence of the proposed framework in capturing and interpreting ontology and CAABAC information. This proficiency is paramount in enabling the GPT to make appropriate access control decisions. For a comprehensive examination of the GPT responses, we introduce two inherently connected key criteria: (1) *Context Comprehension*, representing the system's ability to fully absorb and understand the Ontology and CAABAC information appropriate to the situation at hand, and (2) *Recommendation Effectiveness*, assessing the beneficial nature and practicability of GPT's recommendations. The valuable recommendations generated by the GPT rely on its effective understanding of the contextual information provided. Consequently, a failure in *Context Comprehension* (score below 0.25) immediately results in a zero score in *Recommendation Effectiveness*. We propose a "marking rubric" to assess system responses, mirroring a grading scheme similar to those used for student assignments. This rubric, presented in Table 4.3, allows the evaluation of each question against both criteria, giving scores ranging from 0 to 1. Consequently, a set of 10 questions can achieve a total score ranging between 0 and 10.

TABLE 4.3: Marking rubric for evaluating GPT responses

| Criteria | Potential Scores | Interpretation |
|------------------------------|------------------|--|
| Context Comprehension | 0 - 0.5 | 0: System fails to capture the Ontology and CAABAC information in the evaluated situation. 0.25: System partially captures the Ontology and CAABAC information in the evaluated situation. 0.5: System fully captures the Ontology and CAABAC information in the evaluated situation. |
| Recommendation Effectiveness | 0 - 0.5 | 0: GPT's recommendations are not beneficial, require extensive human improvements, or if Context Comprehension score is 0. 0.25: GPT's recommendations are somewhat beneficial, and require moderate human improvements. 0.5: GPT's recommendations are highly beneficial and require little to no improvements. |

4.5 Evaluations

Post-experiment, we analyzed the data and evaluated the performance of the GPT-Onto-CAABAC framework, providing valuable insight into potential improvements. In the aftermath of our experimental phase, we undertook a rigorous analysis and evaluation of the GPT-Onto-CAABAC framework's performance. This process led to the discovery of invaluable information that could inform potential improvements to the system. Although our evaluation was somewhat speculative due to the lack of real data or observed system behavior, we were able to identify several recurring patterns across all three posts. These patterns included role-specific permissions, policy adherence, patient consent, healthcare purpose, and the need for supervision in certain scenarios.

4.5.1 Scenario Testing with Evaluation Metrics

The GPT-Onto-CAABAC framework was subjected to a series of scenario tests to evaluate its performance. These scenarios were designed to mimic real-world healthcare decisions and the complexities associated with them. The ability of the framework to navigate hospital policies, legal requirements, and dynamic patient-specific contexts was evaluated. The scenarios also tested the framework's adaptability to different roles and their associated permissions. For example, the access rights of a healthcare professional, a relative of the patient or a legal guardian were evaluated under the guidelines of the My Health Records Act 2012 and other similar privacy laws.

In addition to role-based access, the importance of patient consent was also evaluated. The framework was tested for its ability to handle situations where consent could potentially allow individuals who would not typically have permissions, such as friends or siblings, to access the EHR and contact information. Scenario tests also included cases where supervision is required when accessing sensitive data, such as those involving students or interns. The framework's ability to identify and enforce such requirements was evaluated. The framework was also subjected to fault injection testing, where faults or errors were deliberately introduced into the system to test its resilience and robustness. This included scenarios where incorrect or conflicting data was inputted, and the system response was observed. The results of these scenario tests provided valuable insight into the performance of the framework and highlighted areas for potential improvement. The framework demonstrated a high degree of adaptability and robustness, effectively handling a variety of complex scenarios and recovering from introduced faults. However, further testing and refinement are required to optimize the system's performance fully.

4.5.1.1 Scenario Testing

Our rigorous scenario testing, as depicted in Fig. 4.3, provides valuable insight into the performance of the GPT model, particularly in interpreting the legalities associated with EHR access control decisions. We scrutinized the model's interaction with diverse access roles, its application of the My Health Records Act 2012, and its conflict resolution capacity.

Contextual Comprehension: Our evaluation demonstrated the ability of the GPT model to grasp contextual information. In tests where a healthcare professional requested access to a patient's EHR, the model proficiently applied the privacy laws and healthcare protocols embodied in the My Health Records Act 2012. The model discerned that access should be predicated on the professional's role and necessity, appropriately respecting patient privacy. Moreover, in situations involving access by a relative or legal guardian, the model thoughtfully considered the legal and ethical obligations mandated by the Act, demonstrating a detailed understanding of role-based access control.

Recommendations Effectiveness: Our testing highlighted the robust recommendation capabilities of the model. In hypothetical situations involving disagreements over patient consent, the GPT model astutely advised the healthcare professional to seek additional legal or ethical guidance or defer to a higher authority within the organizational hierarchy. This underlines the model's sound comprehension of legal stipulations and its competence in suggesting practical solutions within legal parameters. Nevertheless, it is

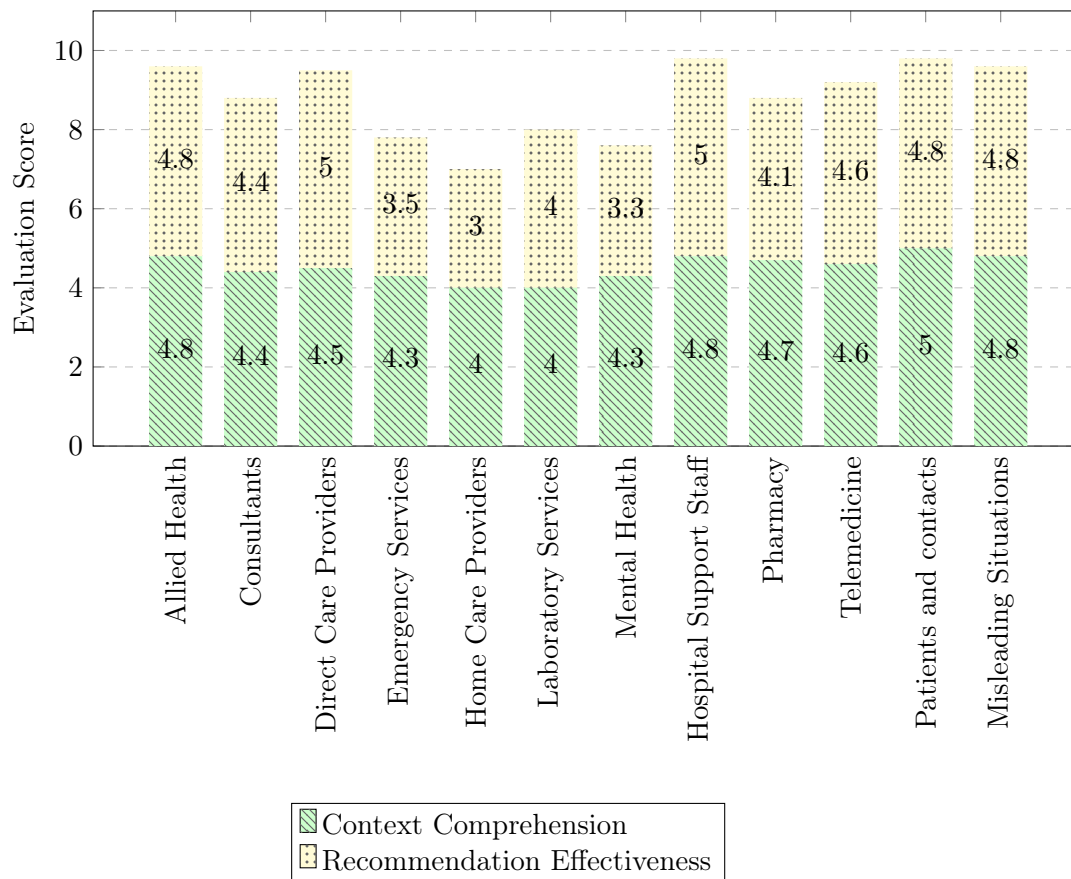


FIGURE 4.3: Evaluation of GPT Answers Per Category (higher is better)

important to recognize that these scenarios were simulated and devoid of real-world conflicts, warranting further investigation of the model's performance under more complex, real-life circumstances.

Overall Performance: The scenario testing indicated a significant potential for the GPT model in legal interpretations related to access control decisions in the EHR. Its apt contextual understanding, combined with practical recommendations, underscores its potential utility as a decision support tool in healthcare. However, while the model showed a comprehensive understanding of the legalities and ethics of EHR access, the need for further evaluations of its conflict resolution capabilities in real-life scenarios remains. Therefore, continuous refinement and testing of the model's capabilities is crucial considering the complexity and evolving nature of healthcare and legal landscapes.

4.5.1.2 Fault Injection Testing

Our GPT-Onto-CAABAC framework, which uses GPT models, ontology systems and context-sensitive attribute-based access control models, plays an instrumental role in resolving dilemmas regarding access to medical data. Its ability to offer a policy-compliant set of options, rather than dictating a single course of action, provides subtle guidance to healthcare professionals. We tested this model by injecting faults in the form of mishaps. These are scenarios in which an individual may mistakenly believe that they have the right to access a patient's EHR due to their personal relationship with the patient. GPT's responses to these fault injection scenarios, in line with the My Health Records Act 2012, were evaluated. In particular, the model exhibited admirable understanding of context and recommendation effectiveness. It demonstrated a clear understanding of the legal boundaries and consistently factored in the nature of the professional's role, their registration as a healthcare provider, and the necessity of patient's consent. GPT responses received high scores in most cases, suggesting a strong alignment with human expectations and interpretation. However, some variances in marking were observed, particularly in scenarios involving close personal relationships, such as spouses or close family members. Although GPT consistently advised the need for patient consent, its recommendations were perceived as slightly lenient considering the intimate relationships involved. This minor deviation could indicate areas where the model's decision-making could benefit from further refinement to handle more complex situations effectively.

Despite these minor inconsistencies, GPT's overall performance in the face of fault injection testing was promising. It demonstrated strong resilience to tackle tricky clinical scenarios, thus underlining its potential as a powerful tool in auditing medical access control risk. The model can help risk auditors identify potential compliance issues and deviations, providing valuable information for training and policy revisions. Our GPT-Onto-CAABAC framework, in essence, represents an innovative intersection of AI and healthcare regulation. Its ability to handle complex, ethically charged situations with respect to access to medical data presents exciting possibilities for the future of healthcare data management. However, as these tests reveal, its usage should be complemented by human oversight to handle complex cases effectively.

4.5.1.3 GPT Responses Patterns

Our GPT-Onto-CAABAC framework, in its interpretation of legal boundaries for EHR access, demonstrates a rich and complex range of responses in different scenarios. These responses, depicted in Fig. 4.4, highlight the multifaceted nature of this AI system and its ability to understand and adapt to sophisticated contexts.

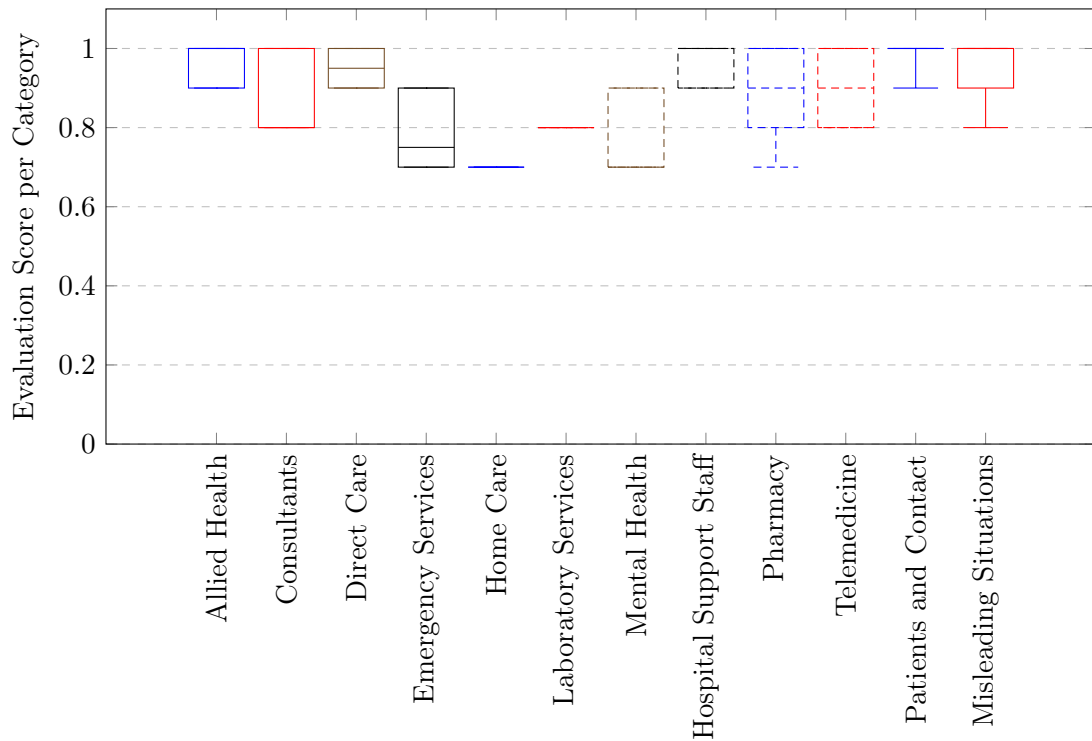


FIGURE 4.4: Variation of Evaluation Scores of GPT Responses By Category

Upon in-depth analysis of the patterns emerging from the GPT's responses, five key categories of variations were identified: role-specific permissions, policy adherence, patient consent, healthcare purpose, and supervision.

- Role-specific Permissions:** As illustrated by the data, role specificity has a profound impact on GPT responses. For categories such as consultants, allied health, and direct care, GPT models showed near-perfect adherence to policy. For roles with less well defined policy boundaries, such as emergency services, mental health, and hospital support staff, a slight decrease in the evaluation score was observed. These lower scores may result from the relative ambiguity in access control policies specific to these roles, requiring a more detailed judgement from the GPT model.
- Policy Adherence:** Policies outlined in the My Health Records Act 2012 form the backbone of access control decisions. The GPT models showed an excellent comprehension of these policies, as observed in high scores in most categories. However, variations exist; in the case of misleading situations or home care, where personal relationships and less formal care settings blur the policy lines, the evaluation scores drop slightly. This may reflect GPT's struggle to balance legal policy with complex human situations.

- **Patient Consent:** Consent is a crucial factor in accessing healthcare data. GPT's interpretation of consent-focused scenarios received creditable scores, especially when dealing with the 'Patients and Contact' category. The slightly lower score in 'Misleading Situations' may be attributed to the ambiguity introduced by the presence of close relationships, which challenges the strict legal interpretation of patient consent.
- **Healthcare Purpose:** GPT's responses accurately reflected the healthcare-centric purpose of access to the EHR, achieving high scores in areas such as direct care, consultants, and telemedicine. Lower scores in home care and emergency services suggest the model's difficulty in perceiving purpose in crisis situations or informal care environments.
- **Supervision:** In situations involving supervised roles, such as students or interns, GPT was adept at incorporating the need for supervision into its responses. The lower score for 'Laboratory Services' may suggest the need for improved model training on subtle roles that might require supervision.

These variations offer valuable insights into the subtle performance of the GPT-Onto-CAABAC framework. The fluctuating scores across categories point to the AI's struggles and successes in interpreting complex legal and ethical issues surrounding EHR access. Although GPT models excel in clearly defined situations, they show difficulty when handling ambiguous or emotionally charged contexts. Therefore, while the GPT model is an impressive tool to interpret access control decisions, these results highlight the essential need for human oversight. Variations in response patterns underscore the ongoing challenge of refining AI models to understand the full complexity of real-life situations and indicate potential areas for future improvement. Interpreting these variations can help develop more accurate and context-sensitive AI systems for the future.

4.5.2 Comparative Evaluation

GPT models such as GPT-3 and GPT-4 have demonstrated notable competencies in understanding and generating human-like text. Their adaptability across various tasks, even without task-specific data, proves beneficial in domains such as healthcare and law, where dynamic interpretations of user roles and corresponding access rights are essential. However, their decision-making process can be time-consuming, contrasting with the immediate decisions rendered by traditional access controls based on pre-set rules and policies. In healthcare, GPT models offer extensive patient histories, suggest relevant medical tests, and help to develop differential diagnoses. Our scenario tests (Subsection

4.5.1.1) demonstrated the adept understanding of the My Health Records Act 2012, effectively handling diverse healthcare roles. However, its efficacy in real-world conflicts requires further exploration. GPT also shows promise in legal contexts, with the ability to interpret complex legal documents, formulate legal arguments, and even predict legal outcomes. Our fault injection tests (Subsection 4.5.1.2) demonstrated that the GPT model provided policy-compliant options even in deceptive scenarios, underscoring its robustness in interpreting legal aspects related to EHR access control decisions.

Traditional access controls, while less adaptable to rule or policy changes and requiring manual adjustments, offer the advantage of speed in decision-making, especially in time-critical real-time scenarios. However, GPT models adapt quickly to new data and context changes, providing a vital edge in settings with evolving access control needs. The extent of this adaptability, for both GPT and traditional models, largely depends on the use-case specifics and system programming. Despite their slower response time, the significant benefits of GPT models lie in their adaptability and flexibility. They are particularly useful for postmortem audits in risk management, employing their capability for detailed text generation to offer valuable insights for risk assessment and mitigation. As revealed by the GPT response patterns (Subsection 4.5.1.3), the variable performance of GPT models under different conditions underscores the need for human oversight and suggests areas for potential improvement.

4.5.3 Ethical and societal implication analysis

In the context of access control of the EHR, ethical and social implications primarily revolve around conflicts that could arise from varying access rights associated with different roles and potential disagreements regarding patient consent. In particular, the scenario tests conducted to evaluate the performance of the GPT-Onto-CAABAC framework did not explicitly present any such conflicts that required resolution. However, potential conflicts could surface in real-world settings. These could be due to contradictions between access permissions of distinct roles, such as healthcare professionals and relatives of the patient, especially when their interests do not align. Similarly, situations may arise where disagreements about patient consent could trigger conflicts, posing a substantial challenge to the decision-making process.

The proficiency of the GPT-Onto-CAABAC framework in addressing and resolving such conflicts can be adequately gauged only when it is confronted with actual conflict scenarios. As such, despite the promising preliminary results from the initial tests, it remains crucial to subject the framework to rigorous and comprehensive testing simulating

real-world conflict scenarios to fully assess its effectiveness and readiness for practical implementation.

4.5.4 Assessment of Transparency and Interpretability

Addressing prevalent concerns around the “black box” phenomenon in AI systems, we made a conscious effort to evaluate the transparency and interpretability of the GPT-Onto-CAABAC framework. The primary objective was to discern whether the framework’s decision-making process and outputs were intuitively understandable and accessible to healthcare professionals or policy makers. The assessment, far from being a superficial overview, entailed a thorough examination of the GPT-Onto-CAABAC framework’s rationale behind EHR access control decisions. This rigorous investigation intended to ensure that healthcare professionals or policy makers could easily understand the logic of the framework, thus facilitating informed decisions regarding access control to the EHR based on the framework’s insights.

Our framework demonstrated consistent response patterns across various scenarios, substantially strengthening its interpretability. Provided satisfactory reasoning based on factors such as role-specific permissions, policy adherence, patient consent, healthcare purpose, and supervision. While processing requests and offering recommendations, it effectively accounted for various aspects defined by the My Health Records Act 2012. The analysis indicated a substantial degree of transparency and interpretability in the framework’s decision-making process, increasing its potential utility in a real-world healthcare setting. Although these promising results are encouraging, continued refinement and testing of the framework’s capabilities, particularly for complex scenarios, are necessary to further enhance its transparency and interpretability. Balancing this need with human oversight, especially in ambiguous or emotionally charged situations, is crucial. The GPT-Onto-CAABAC framework’s transparency and interpretability assessment results demonstrated its ability to offer decision-making processes that are comprehensive, consistent, and accessible to end-users, thereby suggesting its potential as a viable decision-support tool in healthcare settings.

4.6 Discussions

This section examines a comprehensive discussion of the significant issues that emerged during the experiment.

4.6.1 Challenges and Overcoming Strategies

The implementation of the GPT-Onto-CAABAC framework within healthcare care, despite its significant potential, presents several distinct challenges. The complexity of healthcare scenarios, performance and validity issues, and the overarching concern of societal trust necessitate a systematic addressal. However, these challenges also present opportunities for further refinement and innovation.

- **Stability of GPT-generated texts:** In our pilot tests, we found that GPT produces slight variations in its outputs for the same input, primarily linguistic rather than semantic. We propose regular audits and ongoing scrutiny to ensure the consistency and reliability of GPT-generated content. Additionally, implementing feedback loops from end users can provide valuable insights for model fine-tuning.
- **Performance of the GPT models:** With the increasing sophistication and size of GPT models, there is an associated increase in response generation time, making the framework unsuitable for real-time, time-critical decision-making in healthcare. To address this, we recommend continuing performance evaluations and the development of optimization strategies. This may involve parallel processing, model pruning, or exploring hardware acceleration options.
- **Validity of GPT-based decisions:** The potential of GPT models to produce hallucinations, factually incorrect or irrelevant outputs, could lead to non-compliant healthcare decisions. To mitigate this risk, it is crucial to implement continuous validation checks and a verification mechanism[363]. This could involve cross-checking GPT output with trusted resources, implementing peer review mechanisms, or integrating GPT with rule-based systems for sanity checks.
- **Societal trust in AI systems:** The potential for hallucinations and the opaque nature of AI algorithms present a significant challenge in fostering social trust. For this, we advocate strong human oversight, robust mechanisms to monitor the validity of GPT output, and effective public communication strategies. Transparency on model limitations, clear communication about how decisions are made, and maintaining accountability are essential to earning public trust. Additionally, collaboration with regulatory bodies and ethicists to design guidelines and policy frameworks can contribute to social trust.

Addressing these challenges is not a one-time activity, but requires an ongoing cycle of refining and evaluating the GPT-Onto-CAABAC framework. Through continuous iteration, we can improve performance, validate results, improve transparency, and maintain

effective public communication to harness the power of this framework in healthcare decision-making.

4.6.2 Detailed Dataset Discussion

The data set utilized in this study was meticulously selected to ensure the relevance and applicability of our findings to real-world healthcare scenarios. Comprising anonymized EHR data, the data set includes a variety of attributes relevant to healthcare access control decisions, such as patient information, healthcare provider roles, patient-provider relationships, and contextual information such as time of access and location. The data set was derived from a collaboration with a large healthcare provider, ensuring that it reflects the complexity and diversity of real healthcare operations.

To prepare the data set for our experiments, we performed several pre-processing steps. These included anonymization to protect patient privacy, normalization of attribute formats, and categorization of access scenarios into typical healthcare operations (e.g., patient consultation, review of medical records, emergency access). This preprocessing ensures that our experiments accurately reflect the challenges and needs of access control in healthcare settings.

Given the sensitivity and complexity of healthcare data, the selection and preparation of this data set was crucial to the success of our study. Although specific details of the composition of the dataset are proprietary to the collaborating healthcare provider, the described pre-processing steps and the general characteristics of the dataset are provided to aid in the reproducibility of our research findings [51, 125, 364]. Future studies wishing to replicate or extend our work are encouraged to seek partnerships with healthcare providers to secure similar datasets, ensuring that the data used reflect the complexity and nuances of real-world healthcare access control challenges.

4.6.3 Applications in Healthcare Settings

Our GPT-Onto-CAABAC framework offers an adaptable solution that fits a variety of healthcare settings. Its flexibility facilitates its use in healthcare decision-making domains, acting as a proactive recommendation system or as a reactive risk management tool. Traditional security consultations in healthcare care face challenges such as the intensive manual work required to audit complex policies, unclear interpretations of regulations, and the rigidity of adjusting to new policies. These issues, combined with often inadequate information, could affect the effectiveness of consultations. The GPT-Onto-CAABAC framework confronts these challenges head on. LLMs automate auditing,

drastically reducing manual involvement. The natural language skills of the GPT models clarify complex healthcare contexts, and the continuous learning feature of the framework keeps it aligned with changing regulations. This combined prowess offers healthcare professionals a reliable decision-making tool.

Activistically, our framework guides early decision-making stages, presenting policy-aligned alternatives for complex clinical situations. Here, GPT models comprehend detailed patient data, while ontology systems provide context-driven advice based on policy and regulatory interpretations. This cohesive method promotes subtle decision making tailored to each case's specifics. As a reactive mechanism, the GPT-Onto-CAABAC system reviews healthcare decisions after the fact, ensuring that they adhere to legal and organizational standards while highlighting nonconformities. This retrospective review ensures consistent policy adherence, highlights training needs, and pinpoints policy areas that need further clarification. Furthermore, this framework has potential as an educational asset in the training in health care. Through the analysis of previous decisions, it can refine academic syllabi, shedding light on the complex relationship between healthcare care methods, policy mandates, and real patient situations. Despite its obvious value, it remains essential to evaluate the effectiveness of the GPT-Onto-CAABAC framework in diverse healthcare settings, ensuring its continued relevance and contribution to healthcare decision processes.

4.6.4 Expanded use cases beyond EHR

Our GPT-Onto-CAABAC framework has broad applicability across diverse sectors that require complex and detailed access control decisions considering compliance, context and attributes. Here are some potential use cases:

- **Financial Services:** In the financial sector, access controls for sensitive customer data must balance privacy regulations, individual access needs, and security priorities. The framework can aid in compliant access control by considering attributes of financial advisors, the context of customer consent, and privacy laws.
- **Defence organizations:** For defense organizations, granting access to classified data requires strict adherence to security protocols and hierarchies. The framework can incorporate user roles, context-like emergency situations, and classification levels to make informed yet flexible access decisions.
- **Legal Services:** In legal services, client confidentiality is paramount while working with experts in all specializations. The framework can weigh attorney attributes,

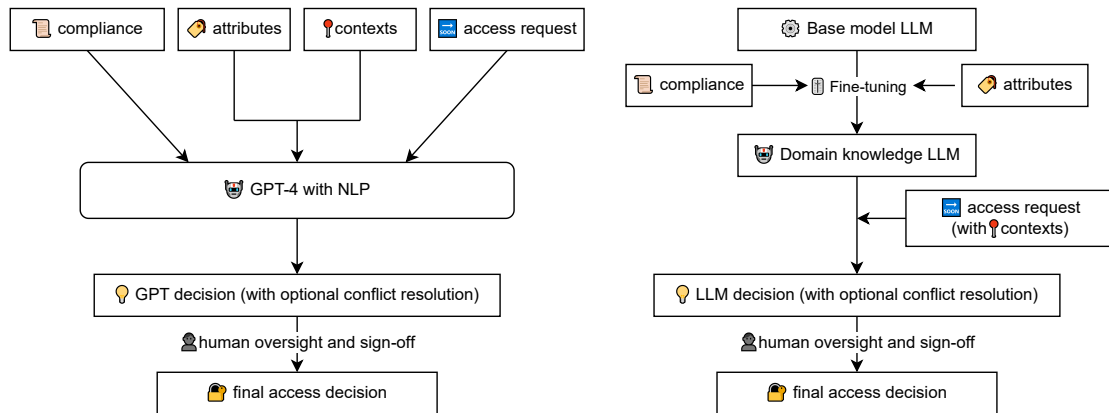


FIGURE 4.5: Comparison of our GPT-4-based prototype (left) and a practical domain knowledge LLM implementation (right)

client permissions, and legal ethics codes to enable secure, yet productive, information sharing.

- **Public Sector:** Government agencies manage huge sensitive citizen data subject to complex regulations. The framework can help navigate user clearances, data types, compliance needs, and transparency laws for responsible public data access.
- **Research Institutions:** Academic research requires collaborations across domains while protecting the privacy of the participants. The framework can balance researcher credentials, study protocols, ethics approvals, and privacy laws to uphold rigorous access control standards.

4.6.5 Translating concept to real-world implementation

While the GPT-Onto-CAABAC framework shows promise as a conceptual model, translating it into large-scale healthcare implementation requires the adoption of a fine-tuned domain knowledge LLM (Fig. 4.5), and requires significant translational research and stakeholder participation. Some key aspects should be considered:

- **Pilot Testing and optimization:** Extensive testing in diverse healthcare settings, institutions, and geographic regions is crucial. This allows for framework optimization and customization based on lessons learned during deployment.
- **Regulatory Approvals:** Securing approvals from healthcare governance bodies and demonstrating compliance are essential prior to full-scale implementation. This ensures that patient safety and security standards are met.

- **Change Management:** Training healthcare professionals on the integration of the framework into workflows is vital. Managing organizational change and addressing adoption barriers smooths the transition.
- **Patient Advocacy:** Incorporating patient perspectives through focus groups and consultation can identify potential ethical concerns early. Their insights further bolster framework transparency.
- **Continuous Improvement:** Updating the framework as healthcare regulations and AI advance is essential. Establishing processes for regular enhancements maintains long-term relevance.
- **Economic Analysis:** Conducting a cost-benefit analysis guides budgeting and resource allocation for development and maintenance. Quantifying the value gained aids in wider adoption.

The Gantt chart, shown in Figure 4.6, visualizes the implementation timeline for 2024. The chart has been derived based on expert estimates and stakeholder inputs.

- **Pilot Testing and optimization** is scheduled for Q1, considering it is the primary phase to validate the framework.
- **Regulatory Approvals** are set in Q2, once preliminary results from pilot tests are available.
- **Change Management** spans from Q2 to Q3, as training and transition management processes often overlap with other tasks.
- **Patient Advocacy** is planned for Q3, ensuring that ethical considerations are reviewed and integrated.
- **Continuous Improvement** begins from Q3 and extends to Q4, emphasizing ongoing updates based on the framework's deployment feedback.
- **Economic Analysis** is conducted in Q4 to guide future resource allocation and budgeting decisions.

This phased translational approach is key to overcoming operational complexities and bridging the gap from conceptual model to field deployment. With diligent pilot testing, stakeholder engagement, iterative improvements, and economic prudence, the GPT-Onto-CAABAC framework can progress from theory to practice.

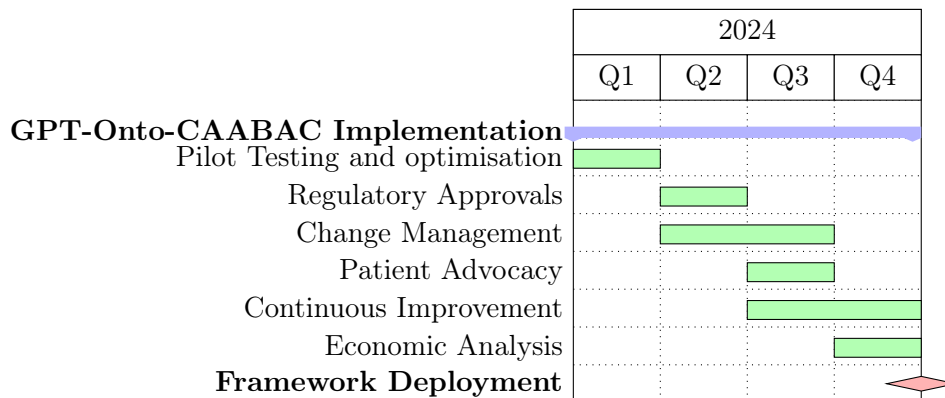


FIGURE 4.6: GPT-Onto-CAABAC Implementation Roadmap for 2024

4.6.6 Expanded Experimental Comparison

In light of the critical importance of robust access control mechanisms in healthcare information systems, this study focused on comparing the GPT-Onto-CAABAC framework against three prevalent access control methods: Role-Based Access Control (RBAC), Attribute-Based Access Control (ABAC), and Context-Aware Access Control (CAAC). These methods were selected due to their wide adoption in healthcare settings and their relevance to the dynamic requirements of Electronic Health Records (EHR) security and privacy. Each of these access control methods offers distinct advantages and limitations, shaping the landscape of EHR security.

RBAC's structured approach simplifies permission management but often lacks the flexibility required for the nuanced access needs of healthcare scenarios [363]. ABAC provides this flexibility by considering a wide range of attributes, yet it can introduce complexity in policy definition and enforcement. CAAC adds contextual decision-making capabilities, offering a more nuanced access control but at the cost of increased system complexity and potential challenges in defining and maintaining context rules.

The GPT-Onto-CAABAC framework aims to address these gaps by integrating the strengths of ABAC and CAAC while leveraging GPT models to interpret and apply complex, evolving access policies. By discussing the limitations and strengths of RBAC, ABAC, and CAAC, we underline the necessity for a solution like GPT-Onto-CAABAC that combines the flexibility of ABAC, the contextual awareness of CAAC, and the advanced interpretative capabilities of GPT models to enhance EHR security and privacy dynamically.

4.6.7 Research limitations

Our research presented in the article focuses primarily on the application of GPT models, ontology systems, and CAABAC models in the context of access control for the EHR. Some potential limitations of our research could include [125]:

- The research might be limited by the quality and quantity of the data used for training the GPT models. If the data are not diverse or comprehensive enough, the models may not perform optimally in real-world scenarios.
- Research may also be limited by the complexity of integrating multiple systems (GPT models, ontology systems, and access control models). This integration might present challenges in terms of system compatibility, data synchronization, and performance optimization.
- Research may be limited by the rapidly evolving nature of both healthcare regulations and AI technologies. The proposed framework might need to be continuously updated to keep up with these changes.

4.6.8 Future research directions

Given the potential limitations of our study, we believe future research could focus on:

- Improving the quality and diversity of training data for GPT models. This could involve collecting more data from a wider range of sources or developing new data augmentation techniques.
- Converting the framework into a domain knowledge LLM tailored for specific use cases, as detailed in Section 4.6.4.
- Exploring more efficient ways to integrate GPT models, ontology systems, and access control models. This could involve developing new algorithms or system architectures.
- Keeping up with the latest developments in healthcare regulations and AI technologies. This could involve regular literature reviews or collaborations with regulatory bodies and AI research institutions.

4.7 Conclusions

Our proposed GPT-Onto-CAABAC framework has advanced EHR access control by incorporating advanced AI capabilities, presenting a dynamic, context-aware model. This integration has the potential to revolutionize healthcare data security and comprehensively address the multifaceted complexities of EHR access control. The ontology-driven component provides a structured methodology for defining crucial concepts such as users, resources, roles, permissions, and contextual data, underpinning coherent access policy articulation, thereby strengthening EHR security. The system adaptability is enhanced through CAAC and ABAC integration, enhancing its applicability across varied healthcare contexts. With the GPT model's inclusion, the system can leverage sophisticated NLP capabilities, facilitating the extraction and interpretation of complex legal and regulatory data, thereby enriching decision-making processes. The design of our model promotes adaptability and efficiency while upholding accountability principles, with inbuilt mechanisms for human evaluation and oversight to foster responsible AI use. Nevertheless, the GPT-Onto-CAABAC deployment is not without challenges. Effective implementation requires substantial resources and expertise, potentially challenging for smaller healthcare entities. Furthermore, given the rapid evolution of healthcare and technology, the model requires regular updates to maintain its relevance. It is essential to balance potential conflicts between the ontology, CAAC and ABAC components and manage the ethical implications of deployment, particularly given the sensitive nature of the EHR data.

Beyond its immediate application in healthcare care, the proposed model shows considerable promise for broader implications. The inherent design of the model showcases immense potential for auditing access control decisions not only in healthcare but across various sectors. Industries with multidimensional policies, rapidly changing contexts, and the need for detailed post-decision audits could significantly benefit from such a model. This opens avenues for the GPT-Onto-CAABAC framework to elevate access control auditing across many critical and dynamic environments. Despite these hurdles and the expanded potential of the model, our GPT-Onto-CAABAC framework represents a significant advance towards integrating state-of-the-art AI capabilities into EHR access control. The dynamism, adaptability, robustness, and context-sensitive attributes of the model enable it to meet evolving healthcare demands while adhering to the prevailing regulations and policies. Future research should focus on optimizing GPT model training data, refining the integration of GPT models, ontology systems, and access control models, and staying abreast of healthcare regulations and AI technologies. As the field progresses, we anticipate that the GPT-Onto-CAABAC model will continue to be a

novel and adaptable solution, enhancing its efficacy in diverse healthcare scenarios, and pushing the boundaries of the application of AI in healthcare.

Abbreviations

| | |
|--------|--|
| ABAC | Attribute-Based Access Control |
| AI | Artificial Intelligence |
| CAABAC | Context-Aware Attribute-Based Access Control |
| CAAC | Context-Aware Access Control |
| EHR | Electronic Health Record |
| GPT | Generative Pre-trained Transformers |
| LLM | Large Language Model |
| NLP | Natural Language Processing |
| RBAC | Role-Based Access Control |

Declaration of Co-Authorship - Chapter 4 PART B



THE NEW WAY TO DO UNI

OFFICE FOR RESEARCH TRAINING, QUALITY AND INTEGRITY

DECLARATION OF CO-AUTHORSHIP AND CO-CONTRIBUTION: PAPERS INCORPORATED IN THESIS


This declaration is to be completed for each conjointly authored publication and placed at the beginning of the thesis chapter in which the publication appears.

1. PUBLICATION DETAILS (to be completed by the candidate)

| | | | |
|------------------------------|---|-------------------------------|--------------|
| Title of Paper/Journal/Book: | Enhancing Health Information Systems Security: An Ontology Model Approach | | |
| Surname: | Nowrozy | First name: | Raza |
| Institute: | Institute for Sustainable Industries and Liveat | Candidate's Contribution (%): | 80 |
| Status: | | | |
| Accepted and in press: | <input type="checkbox"/> | Date: | |
| Published: | <input checked="" type="checkbox"/> | Date: | 11 July 2023 |

2. CANDIDATE DECLARATION

I declare that the publication above meets the requirements to be included in the thesis as outlined in the HDR Policy and related Procedures – policy.vu.edu.au.

| | |
|---|------------|
|  | 06-02-2024 |
| Signature | Date |

3. CO-AUTHOR(S) DECLARATION

In the case of the above publication, the following authors contributed to the work as follows:

The undersigned certify that:


1. They meet criteria for authorship in that they have participated in the conception, execution or interpretation of at least that part of the publication in their field of expertise;
2. They take public responsibility for their part of the publication, except for the responsible author who accepts overall responsibility for the publication;



**VICTORIA
UNIVERSITY**

THE NEW WAY TO DO UNI

3. There are no other authors of the publication according to these criteria;
4. Potential conflicts of interest have been disclosed to a) granting bodies, b) the editor or publisher of journals or other publications, and c) the head of the responsible academic unit; and
5. The original data will be held for at least five years from the date indicated below and is stored at the following **location(s)**:

| Name(s) of Co-Author(s) | Contribution (%) | Nature of Contribution | Signature | Date |
|-------------------------|------------------|---------------------------------|---|-----------|
| Khandakar Ahmed | 20% | Methodology, Review, Proof Read |  | 7/02/2024 |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |

Updated: September 2019

PART B: ENHANCING HEALTH INFORMATION SYSTEM SECURITY: AN ONTOLOGY MODEL APPROACH

NOTE: The content of this chapter has been accepted and published by *Health Information System Conference, Melbourne*

Nowrozy, R., et al. (2023, July). Enhancing Health Information Systems Security: An Ontology Model Approach, *Health Information System Conference, Melbourne, 2023*. (Accepted and published)

4.8 Introduction

Health Information Systems (HIS) are vital for healthcare organizations to securely manage sensitive patient data. However, they are also susceptible to cyber attacks that can compromise patient data confidentiality, integrity, and accessibility [363, 365, 366]. The widespread implementation of Electronic Health Records (EHR) in modern healthcare systems has amplified these concerns [367–369], leading to an urgent need to reevaluate and strengthen the security and privacy of sensitive patient data.

This article delves into the current state of information security associated with EHR, pinpointing the gaps and limitations in contemporary research and practice [370–372]. By exploring the existing state of EHR security research, it becomes evident that there are significant shortcomings in the current approaches to protecting these vital systems [373, 374]. To address these challenges, we introduce a novel and comprehensive security ontology model specifically tailored to enhance the protection of health information systems [375, 376].

To address this concern, [377–379] propose a Security Ontology Model for HIS that can help organize and capture security concepts and their relationships. This model, central to our study, offers a standardized approach to HIS security, aiming to provide a more robust and holistic understanding of security aspects in the context of EHR systems [380, 381]. The proposed model integrates multiple security aspects, such as

confidentiality, integrity, and availability, ensuring a comprehensive protective framework for health information systems [382, 383].

At the core of our proposed model lies the conceptual ontology security model, providing a formal representation of security concepts and their interrelationships within EHR systems [384–386]. This model encompasses various elements, including Health Information, EHR Security Conditions, and SWRL (Semantic Web Rule Language) rules [387–389]. By integrating a diverse range of access control strategies, such as Role-Based Access Control (RBAC), Attribute-Based Access Control (ABAC), and Mandatory Access Control (MAC), the model offers the flexibility and adaptability needed to cater to different organizational security requirements [6, 390, 391].

Combining these access control strategies with SWRL rule bases and security policy ontology, our model facilitates fine-grained control over access to health information [72, 392, 393]. This not only ensures adherence to relevant regulations and policies but also significantly enhances the security posture of HIS [394–396].

Despite the considerable contributions of this research, we acknowledge the presence of certain limitations, including the need for further validation and testing in real-world environments [397–399]. The article concludes by underscoring future research challenges, such as the development of more sophisticated access control mechanisms and the effective integration of the proposed model with existing health information systems [400–404].

4.8.1 The current status of information security related to EHR

Because it ensures the confidentiality, integrity, and availability of sensitive patient information, information security is an essential component of “*electronic health records systems*”. However, the increasing number of data breaches and security breaches in the healthcare industry raises concerns about the current state of EHR information security. The absence of a comprehensive and standardized ontology model that can effectively capture and represent the various concepts and security relationships in the domain of health information systems is one of the major obstacles to the security of EHRs. Ontology is a formal representation of knowledge that identifies concepts, relationships, and rules in a domain [405]. An ontology model can provide a structured and organized framework for understanding, managing, and mitigating security risks and threats in the context of EHR security. There is a lack of consistency and fragmentation in current approaches to EHR security, making it difficult to understand and apply security measures. In addition, a robust and adaptable ontology model that is able to effectively capture

the ever-evolving security requirements and technologies is necessary due to the complex and dynamic nature of health information systems.

A comprehensive security ontology model for health information systems is required to address these difficulties. Authentication, authorization, encryption, audit logs, data integrity, and data privacy should all be included in this model, along with their connections to EHR components such as users, applications, data repositories, and communication channels. The “*HIPPA*” and the “*National Institute of Standards and Technology (NIST)*” Cyber Security Framework are two examples of relevant healthcare-specific security standards, regulations, and best practices. The security metaphysics model should be founded on a strong hypothetical establishment, like the brought together “*Displaying Language*”, and should be designed to be versatile, extensible, and interoperable with other principles and frameworks of medical services. In addition, it should be able to adapt to a variety of EHR environments, such as cloud-based, mobile, and interoperable EHR systems. The absence of a comprehensive and standardized ontology model presents significant obstacles to the current state of EHR information security. To effectively manage and mitigate security risks and threats in EHR systems, it is essential to develop a model of health information systems security ontology that is robust and adaptable [406]. This will empower medical services associations to carry out steady and viable safety efforts and protect delicate patient data from unapproved access, information breaks, and other security events.

4.8.2 The current state of research and a brief introduction of its inadequacies of security EHR

The growing recognition of the importance of information security in healthcare is reflected in the steady increase in research on security in “*electronic health record systems*”. Regardless of the progress made, there are still deficiencies in the current status of the examination in the security of the EHR. The absence of a comprehensive and standardized approach to security in EHR systems is one of the main deficiencies. The holistic nature of information security is ignored in the majority of current research studies, which instead concentrate on specific aspects of EHR security, such as authentication or encryption. Due to this fragmented approach, security measures may be insufficient or inconsistent, leaving other parts of the EHR system vulnerable. For effective protection of patient information, a comprehensive strategy is necessary that takes into account all aspects of EHR security—including technical, organizational, and human factors—is necessary. Another deficiency is the restricted reconciliation of safety guidelines and best practices in EHR research. HIPAA and the NIST Cybersecurity Framework are two established security standards and regulations in the healthcare industry, but they are

not always incorporated into EHR research studies [407]. This could make it difficult to put the research findings into practice because they might not be in accordance with the actual security requirements. Implementing effective security measures in real-world EHR systems can be made easier by incorporating relevant security standards and best practices into EHR research.

Moreover, the unique idea of EHR frameworks and the developing danger scene require persistent exploration and updates to address emerging security challenges. However, some studies may not be up-to-date and do not keep up with the rapidly evolving technology and threat landscape in the healthcare industry. Effective security measures that can withstand the constantly evolving threats in EHR systems can only be developed by staying up-to-date with the latest security threats and technologies. In addition, more empirical research is required to assess the efficiency of security measures in actual EHR systems. There are a lot of theoretical frameworks and concepts for EHR security that have been proposed in research studies, but there are not enough empirical studies to prove that these measures work in the real world [407]. Practical evaluations of security measures in actual EHR systems can be included in empirical research, which can provide valuable insights into their effectiveness and identify potential areas for improvement.

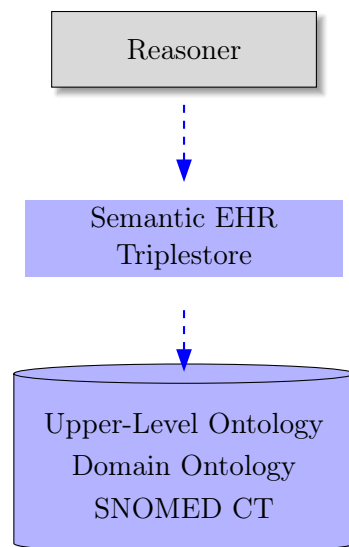


FIGURE 4.7: Knowledge structure of EHR triplestore.

4.8.3 The Contributions and the Outline of the Study

The creation of a security ontology model for HIS is the main contribution of this chapter. There are three layers in the model: the conceptual layer, the implementation layer, and the foundational layer. The central layer provides many fundamental security ideas, including secrecy, respectability, accessibility, confirmation, and approval. The

conceptual layer introduces new security concepts such as risk, threat, vulnerability, and countermeasure [405]. It also defines the relationships that exist between the various security concepts. In order to guarantee the safety of HIS, concrete examples of security measures are provided in the implementation layer. There are a number of advantages to the security ontology model described in this chapter. To begin with, it provides healthcare organizations with a comprehensive understanding of security concepts and their relationships, which can help them develop efficient security strategies. Second, it can be used as a foundation for HIS security standards and guidelines. Third, it may make it easier for various stakeholders involved in the creation and implementation of HIS security measures to communicate with each other and work together.

TABLE 4.4: MEDICATION ADMINISTRATION PATTERN TRIPLE-BASED REPRESENTATION

| # | Subject | Predicate | Object |
|----|------------------------------|--------------|-------------------------------|
| 1 | btl:Plan | isRealizedBy | sct:Medication Administration |
| 2 | sct:MedicationAdministration | hasFocusOn | sct:Pharmaceutical Product |
| 3 | sct:MedicationAdministration | hasRoute | sct:RouteOf Administration |
| 4 | sct:MedicationAdministration | hasStartTime | btl:PointInTime |
| 5 | sct:MedicationAdministration | hasEndTime | btl:PointInTime |
| 6 | sct:MedicationAdministration | hasDuration | btl:Duration |
| 7 | sct:MedicationAdministration | hasFrequency | shn:Frequency |
| 8 | sct:PharmaceuticalProduct | hasComponent | sct:Substance |
| 9 | sct:PharmaceuticalProduct | hasDose | shn:PhysicalQuantity |
| 10 | sct:PharmaceuticalProduct | hasForm | shn:DrugDoseForm |
| 11 | sct:Substance | hasStrength | shn:PhysicalQuantity |
| 12 | sct:Substance | hasForm | shn:DrugDoseForm |
| 13 | shn:PhysicalQuantity | hasValue | xml:double |
| 14 | shn:PhysicalQuantity | hasUnits | shn:MeasurementUnits |

This above tables, represents the contribution of security ontology-based health information system in Electronic Health Information System based administration. To guarantee the security and scalability of “Health Information Systems”, this chapter investigates the creation of a security ontology model and the obstacles that must be overcome. An introduction to HIS and the importance of data security in healthcare is provided at the beginning of the study. The need for a security ontology model to address HIS’s complex security challenges of HIS is discussed [407]. The first section of the study focuses on the difficulty of creating a comprehensive security ontology model due to the complexity of HIS and the lack of standardization. Additionally, in this section, the importance of developing a consistent strategy to ensure that patient data remain safe and accessible is highlighted. The second part of the study investigates the advancing danger scene examined by HIS and the inadequate security principles. Talk about the need to create

a security ontology model that can deal with changing threats and keep the system's security current. The third section of the study discusses the obstacles to the creation of a security ontology model, including user acceptance and restricted data access. Examine the importance of ensuring that the system meets user requirements and the requirement for user-friendly interfaces that can be easily integrated into existing systems. The final section of the study is devoted to scalability, a significant obstacle that HIS must overcome as the volume of patient data continues to increase [406]. It discusses the necessity of creating a security ontology model that can scale to meet the rising demand for data storage without jeopardizing the security of the system. The study emphasizes the need to address these issues to create a comprehensive and efficient HIS security ontology model. Developers can do this to ensure that patient data remains safe and accessible even as data volumes continue to increase.

4.9 Research Motivations

4.9.1 Motivating EHR Use Case Scenarios

Providing clinicians with immediate access to patient data and facilitate seamless communication between healthcare providers, "eHealth Records" are an essential component of modern healthcare delivery. However, strict security measures are required to safeguard patient privacy due to the sensitive nature of health information. Despite this, inadequate access control policies for electronic health records (EHRs) have been found to be the cause of data breaches and potential harm. The sharing of patient data between various healthcare providers is one use case scenario that highlights the need for a new security model and ontology for health information systems. A safe and effective method of sharing patient medical information is required when they are referred to specialists or require care from multiple providers. However, current access control policies have a tendency to be too restrictive, making it difficult for healthcare providers to gain access to the data they require to provide the appropriate care. The use of mobile health applications is another scenario in which a new security model or ontology is required [408]. A robust security framework protects patient information while still allowing easy access to health data is required as more patients use mobile apps to manage their health. Insufficient access control policies for mobile health apps can result in data breaches for patients.

4.9.2 Scenario 1: Patient-Centered Care Coordination

Patients in today's healthcare system frequently receive care from multiple providers, which can result in inefficient and fragmented care. Electronic health record frameworks can further develop care coordination by allowing various providers to access and share patient data. However, this also raises concerns about the security and privacy of patients. A new security model/ontology with access control policies is required to address these issues. The model should make it possible for healthcare providers to seamlessly collaborate, while also ensuring that only authorized individuals have access to patient information. Patient-centered care coordination is one possible use case for this model. A patient with a complicated medical history is being treated by multiple healthcare providers in this scenario [409]. The patient's medical history, medications, and test results are included in the EHR. Policies that restrict access to this information to only authorized providers involved in patient care should be included in the security model to guarantee both the privacy and the security of the patient. Additionally, the model should allow these providers to securely share information in order to ensure efficient and coordinated care.

4.9.3 Scenario 2: Patient Data Privacy

There have been a number of data breaches in healthcare systems in recent years, resulting in the loss of sensitive patient data. The likelihood of data breaches increasing as more medical facilities use "electronic health records" to manage patient information. This scenario clarifies the need for a new security model or ontology that can better protect the privacy of patient data. Consider the instance of an intriguing clinical patient condition that they would rather not be freely known. This patient trusts their healthcare provider with the confidentiality of his medical records. However, a nonauthorized individual gains access to the electronic health record (EHR) system and can view the patient's medical record, which includes details about his rare medical condition. The patient could suffer severe consequences as a result of this breach of privacy, such as loss of employment or social exclusion [410]. A new security model or ontology that provides more granular control over access to patient data is required to prevent such privacy breaches. The principle of least privilege should serve as the foundation for access controls, ensuring that only individuals who require particular patient data are granted access. Additionally, only a limited amount of access should be granted, and it should be regularly audited to catch any attempts at unauthorized access. These measures can help patients trust the healthcare system and prevent data breaches from causing harm.

4.9.4 Scenario 3: Emergency Room Admission

Consider a situation in which a patient is admitted to the emergency room after a car accident. The patient is unconscious and unable to provide any information about his medical history. In this case, nurses and doctors will need to quickly access the patient's EHR to learn about the patient's medical history, allergies, medications, and any conditions they may already have. However, there are a number of security risks associated with accessing the patient's EHR. Data breaches, unauthorized access, and tampering could all occur if the patient's electronic health record (EHR) is not safeguarded by a robust security model. Additionally, if the EHR is not accessible promptly, the medical team may not be able to provide the patient with timely treatment. As a result, it is essential to have an effective and safe access control policy in place in the event of an emergency like this [411]. The patient's electronic health record (EHR) must be accessible by only authorized personnel, and this access must be quick and efficient, according to the access control policy. In emergency situations where time is of the essence, this is especially crucial. Access control policies that prioritize quick and secure access to patient EHRs in emergency situations while maintaining patient privacy and security must be included in a new security model or ontology for health information systems.

4.9.5 Scenario 4: Clinical Research

The purpose of clinical research is to improve patient care and advance medical knowledge through the collection, analysis, and interpretation of data. Because they enable widespread access to patient data, electronic health records are an essential resource for clinical research. However, there are a few difficulties in getting to this information, including security concerns and administrative prerequisites. In this scenario, a group of researchers is looking to see how well a new treatment for a rare disease works. To identify potential study participants and collect medical history data, researchers need access to electronic health records. However, researchers must obtain permission from patients or their legal representatives because access to these data is restricted due to privacy concerns.

It is possible that the requirements of clinical research cannot be met by the access control policies that are currently in place for health information systems. For researchers to carry out their studies, they need access to a large amount of data. However, to protect the privacy of patients, these data must be secured. In addition, compliance with regulations is necessary to guarantee the ethical conduct of clinical research [412]. Access to patient data for clinical research could be made easier and safer with a new security model or ontology and updated access control policies for health information

systems. While ensuring patient privacy and adhering to regulatory requirements, this would facilitate the development of new treatments and the advancement of medical knowledge.

4.9.6 Scenario 5: Unauthorized Access and Disclosure of Mental Health Records

Information about a person's mental health history, including their diagnosis, treatment, and medications, can be found in mental health records. These data are exceptionally private and should be protected to prevent unapproved access or disclosure. However, several factors, including inadequate access controls, social engineering attacks, or malicious insiders, can result in unauthorized access to or disclosure of mental health records in some instances. For example, imagine a situation where a patient's psychological well-being records are gotten to by an unapproved individual. This could happen if a hacker gets into the hospital's information system, a malicious employee gets into the records, or a healthcare provider accidentally gives the information to a stranger [413]. Discrimination, stigma, and a breach of trust can all result from unauthorized access or disclosure of a patient's mental health records. In this way, there is a requirement for another security model/metaphysics for well being data frameworks that can address the particular difficulties of safeguarding emotional well being records. Access control policies for mental health records should be included in this model to prevent unauthorized access and disclosure. In addition, the model should take into account the ethical and legal repercussions of disclosing information about mental health and guarantee that the patient's rights to privacy and confidentiality are safeguarded.

4.9.7 Scenario Analysis

Health information systems must have access control policies and a robust security model in place, as shown by the scenarios above. The analysis of these scenarios reveals several common themes about the potential dangers of access to unauthorized patient data. First, the scenarios highlight the potential for identity theft, insurance fraud, and even blackmail as a result of data breaches and the theft of sensitive health information. Second, the scenarios demonstrate the importance of access controls to ensure that patient information is accessible only to authorized personnel. Inadequate authentication protocols or weaknesses in access control policies, such as sharing passwords, facilitated unauthorized access to patient data in a number of instances. Third, the scenarios emphasize the need for auditing and monitoring capabilities to keep track of who and when

accesses patient data [414]. In some instances, the unauthorized access was only discovered after the fact, making it difficult to locate the offender and repair the damage. Fourth, the scenarios show how important it is to train staff on the best ways to protect patient information and to have clear policies and procedures in place to deal with security incidents and breaches. The scenarios highlight the need for a new security model/ontology and access control policies that can better address the risks associated with unauthorized access to patient information in healthcare systems.

4.10 Conceptual Ontology Security Model

4.10.1 Background

As the use of “electronic health records” grows in the healthcare sector, protecting the privacy and security of patient health information has become an important concern. Identity theft, insurance fraud, and medical malpractice are just a few of the serious consequences that can result from unauthorized access to EHRs. To maintain patient privacy and prevent such incidents, access control policies are necessary. However, it can be difficult and complicated to create efficient access control policies. A well-organized strategy for developing and carrying out these kinds of policy can be provided by a conceptual ontology security model. It helps to define access control policies for each stakeholder by identifying stakeholder, health information, and the relationships between them [415]. Patient health information can be protected from unauthorized or malicious access through a well-designed security model, which ensures that only authorized personnel have access to it.

4.10.2 Ontology Conceptual Security Model

Health information system access control policies can be designed using the conceptual security model of the ontology. By separating the framework into elements, traits, and tasks, the model considers a reasonable comprehension of what should be secured and how. In addition, it provides a structured method for defining access control policies that can be applied uniformly throughout the system. Using this model can help healthcare organizations ensure that authorized personnel have access to patient information, which is essential for protecting patient privacy and preventing data breaches. The model enables fine-grained control over who can access sensitive data by specifying which entities can perform which operations on which attributes [419]. A foundation for auditing system access is also provided by the conceptual ontology security model. The system is

TABLE 4.5: Data Quality Ontology – Key Concepts in HER Applications

| Concept | Definition | References |
|---------------------|---|---------------------|
| Measure | An aspect of data quality that quantifies a characteristic of the data. | [405–407, 416, 417] |
| CorrectnessMeasure | Measures that assess whether the data that exist in the dataset is true. | [406, 407, 416–418] |
| ConsistencyMeasure | Measures that assess data conformance to constraints, rules, and restrictions of the domain. | [405–407, 416, 417] |
| CompletenessMeasure | Measures that assess whether a truth about the world is contained in the data. | [405–407, 416, 417] |
| CurrencyMeasure | Measures that assess the timeliness of the data to represent the Domain and Task. | [405–407, 416, 417] |
| MeasurementMethod | A series of steps used to quantify an aspect of data quality for a measure. | [405–407, 416, 417] |
| Measurement | The process of performing a MeasurementMethod to produce a measurement result | [405–407, 416, 417] |
| MeasurementResult | The quantity produced by a Measurement Method. | [405–407, 416, 417] |
| Metric | Statistics for how a measurement result varies over time or other dimensions. | [405–407, 416, 417] |
| Dataset | The entire set of representations that are being assessed. | [405–407, 416, 417] |
| Representation | The lowest level, atomic piece of information that exists in the data being evaluated (also known as a data field, observation, value). | [405–407, 416, 417] |
| DomainConcept | Concepts in the clinical Domain and Task of interest that are mapped to representations in the set of data being assessed. | [405–407, 416, 417] |
| Domain | A separate ontology describing the clinical domain of interest. | [405–407, 416, 417] |
| Task | A separate ontology describing the specific purpose of using the data. | [405–407, 416, 417] |

able to identify any attempts at unauthorized access and take the necessary action by keeping track of which entities carried out which operations on which attributes. When it comes to creating access control policies for health information systems, the conceptual security model of ontology is a useful tool. By providing an organized way to deal with security plans, it can help medical care associations keep up with patient protection and prevent information interruptions.

4.10.3 Identifying Stakeholders

A crucial step in developing EHR access control policies is identifying stakeholders. To protect patient privacy and data security, access to patient records must be carefully controlled for each stakeholder, each of whom plays a distinct role in the healthcare sector. Because they have the right to access their own health records, patients are one of the most significant stakeholders. They should be able to read and make changes to their own records, but they should not be able to access the records of other patients or sensitive information that isn't related to their care. Access to patient records is required for treatment by healthcare providers, including doctors, nurses and other staff members [420]. Depending on your role and responsibilities, your permissions can vary. For example, nurses may only be allowed to read patient records, while physicians may be permitted to read, write, and update patient records. In addition, it is essential to ensure that healthcare providers have access only to patient records that are relevant to their duties. Guarantors and controllers are different partners who might expect admission to patient records for specific purposes. To process insurance claims, insurers may need access to patient records, and regulators may need access to patient records to verify that healthcare regulations are being followed. To ensure that they only have access to the data they need to carry out their responsibilities and to safeguard patient privacy, their permissions should be tightly controlled. In general, effective access control policies for EHRs can only be created by determining the roles and permissions of stakeholders.

4.10.4 Identifying Health Information

Health information is any information about a patient's medical history, diagnoses, treatments, medications, and test results in the context of the healthcare industry. Because it is personal and sensitive, this information needs to be protected to protect the patient's privacy and prevent unauthorized access. When creating access control policies, it is essential to determine the types of health information that need to be protected. Only healthcare providers who are directly involved in patient care and have the patient's permission can access sensitive health information, such as mental health records or HIV status, according to policies. Moreover, approaches can be characterized to guarantee the privacy of patient data during transmission and capacity. Encryption can be used to protect patient data from unauthorized access and interception. Furthermore, policies can be defined to guarantee the accuracy and completeness of patient records—the integrity of health information [421]. By restricting write and delete operations to authorized personnel, access control policies can be designed to prevent unauthorized modifications or deletions of health information. The accuracy of patient records and the prevention of

medical errors can both benefit from this. To ensure the privacy and security of patient information in health information systems, it is essential to determine the types of health information that need to be protected and to create access control policies based on the sensitivity and significance of that information.

4.11 Security or Policy Ontology

To effectively manage patient health information, the healthcare industry has increasingly relied on electronic health records (EHR). Nevertheless, safeguarding the protection and security of these sensitive data presents critical difficulties. As a result, access control policies are absolutely necessary to safeguard patient privacy and prevent unauthorized access to EHRs. A structured approach to designing such policies can be provided by a security or policy ontology to accomplish this objective. The privacy ontology that was previously created will be expanded in this section to include security conditions and policies for the various stakeholders in the healthcare industry. Access control policies for various entities, attributes, and operations will be defined using the formal policy model (E, A, O). Partners expecting admittance to well-being data will also be recognized, and security conditions will be characterized for different kinds of well-being data. Access control policies for various stakeholders in the healthcare industry can be structured by developing a security or policy ontology. By preventing unauthorized access to patient health records, this will help protect their privacy [422]. In addition, the ontology will define security conditions for various types of health information to ensure that only authorized individuals can access sensitive data, such as HIV status or mental health records. Policies can be enforced using SWRL rules to ensure that healthcare providers only have access to patient records that are relevant to their care and that insurers only have access to records needed to process insurance claims. The security/policy ontology offers a methodical approach to the creation of access control policies that can help preserve the privacy of patients and guarantee the safety of their medical records.

4.11.1 Security or Policy Ontology

A structured approach to designing access control policies for stakeholders in the healthcare industry is provided by the security or policy ontology, which builds on the privacy ontology. Access control policies are defined by its entities, attributes, and operations. The ontology is represented by the notation (E, A, O), where E stands for the set of entities, A for the set of attributes, and O for the set of operations. Patients, healthcare providers, insurers, regulators, and researchers are all entities in the security or policy

ontology [423]. These entities have roles, permissions, and locations as attributes. Activities address the activities that can be performed on these elements, such as reading, writing, and executing.

A formal policy model that serves as a framework for defining access control policies is the security or policy ontology. It lets us specify which entities have the ability to carry out which operations on which attributes. A policy might say, for example, that nurses can only read patient records, while doctors can read and write them. Access to sensitive health information, such as HIV status or mental health records, can also be restricted by policies. The security or policy ontology offers a comprehensive strategy to manage health information by expanding the privacy ontology. It protects patient privacy and ensures confidentiality, integrity, and accessibility of healthcare information [424]. The cosmology can be utilized to configure access control approaches that address the issues of various partners and guarantee consistence with medical services guidelines.

4.11.2 Entities

In the policy or security ontology, entities are the objects that need to be protected. Patients, healthcare providers, insurers, regulators, and researchers are the entities in the healthcare industry. For access to EHRs, each entity has its own set of roles and permissions.

1. Patients

Patients have the right to access their own health information because they are the primary stakeholders in the healthcare sector. Depending on your preferences and the laws and regulations of the nation in which you live, patients may have different levels of access to their health information. It is possible that patients will be able to view their health information, make changes to specific fields, or grant access to particular healthcare providers.

2. Providers of Healthcare

Access to patient health information is required for treatment by healthcare providers, including doctors, nurses, and other medical personnel. Depending on their roles and responsibilities, healthcare providers may have different levels of access to patient health information [425]. For example, a primary care physician might have full access to a patient's medical history and the results of tests, while a specialist might only have access to parts of the patient's health record that are relevant to their area of expertise.

3. Insurers

To process insurance claims, insurers need access to patient health information. To verify the legitimacy of the claim and determine the appropriate amount of coverage, they may require access to specific parts of the patient's health record. It is possible that insurers will not be able to see all of a patient's sensitive data and may have limited access to health information.

4. Regulators

In order to carry out their responsibilities, regulators need access to patient health information. They also monitor compliance with healthcare care regulations. Although regulators may have access to a variety of patient health information, they are obligated to maintain strict confidentiality.

5. Researchers

For research purposes, researchers may need access to patient health information. They must adhere to strict confidentiality guidelines and may need access to specific parts of the patient's health record for their research [426]. In order to define access control policies and ensure that patient health information is protected from unauthorized access, it is essential to identify the entities in the security or policy ontology.

4.11.3 Attributes

The attributes of the system's entities are used to define their characteristics in the security/policy ontology. Roles, permissions, and locations are among these attributes. The responsibilities and privileges of the various system entities are defined by roles. For example, healthcare providers play the role of "provider" for patients they treat, while patients play the role of "owner" for their health records. The kinds of operation that entities are allowed to carry out on the EHRs are determined by these roles. Consents are used to indicate the activities that a substance is permitted to perform on the EHRs. For example, a patient might be allowed to peruse and refresh their own well-being records, while a medical care supplier might be permitted to peruse, compose, and update the records of their patients. Consents are characterized in view of the duties of substances and the types of EHR that they need to reach [427]. The physical locations from which entities can access the EHRs are specified by locations. For example, a medical care supplier may be allowed to access EHRs just from their work environment, while a patient may be allowed to access their own well-being records from any area. Because it helps to prevent unauthorized access from outside the healthcare organization, this quality is especially crucial to ensuring the security of electronic health records (EHR).

Time-based access, which restricts an entity's access to the EHRs for a predetermined amount of time, and purpose-based access, which restricts an entity's access based on the purpose for which the EHRs are being accessed, are two additional attributes that can be used in the security/policy ontology. The EHRs are only accessed by authorized parties and for legitimate purposes due to these characteristics. The ascribes to the security / strategy metaphysics characterize the qualities of the elements in the framework and decide the types of activities they are allowed to perform on the EHR [428]. They are necessary to safeguard the privacy and security of EHRs and to prevent unauthorized access by parties who do not have a legitimate need to do so.

4.11.4 Operations

For defining access control policies for various stakeholders, the operations in the security or policy ontology are crucial. The permissions granted to these operations can vary depending on the stakeholder's role and responsibilities. They specify the actions that can be performed on entities. Read, Write, and Execute are the three most commonly performed operations in healthcare. A user can view or access information from a record using the read operation. For example, a healthcare professional may be required to read a patient's medical history to provide the appropriate treatment. In a similar vein, a researcher might require access to particular data in order to carry out an investigation or analysis. A user can modify or update the information in a record using the write operation. After an appointment, for example, a healthcare provider may be required to update a patient's diagnosis or prescription. However, not everyone should be able to change information without permission [429]. For example, a patient should be able to update their personal information, but not their medical history. A user can carry out a specific action on an entity thanks to the execution of the operation. For example, a physician may require a diagnostic test. In a similar vein, a researcher might be required to execute a particular algorithm on particular data to carry out an analysis. By defining which entities are permitted to carry out which operations on which attributes, access control policies can be established. A policy might say, for example, that nurses can only read patient records while doctors can read and write them. In a similar vein, a policy may allow a patient to read and update their personal information, but only healthcare providers can do so. These policies provide appropriate access to essential information for healthcare providers, insurers, regulators, and researchers while preserving patient privacy and confidentiality.

4.11.5 Security Conditions

As a framework for defining access control policies, security conditions are an essential component of the security or policy ontology. The entities that are permitted to carry out operations on which attributes are defined by these policies. The security conditions help maintain the privacy and confidentiality of patient information and are based on the sensitivity and significance of health information. The policy of “role-based access control” is one of the defining security conditions. The user’s role determines who has access to EHRs, according to this policy. For example, healthcare providers might play various roles such as doctor, attendant, or professional, and every job might have various consents to enter patient records. Patients may be able to view and update their own records with the owner role. This RBAC policy helps to ensure that sensitive health information is accessible only to authorized users. The “attribute-based access control” policy is another security condition that can be defined. Access to EHRs is determined by user characteristics and patient records, according to this policy [430]. A policy might say, for instance, that sensitive health information, such as a patient’s HIV status or mental health records, can only be accessed by healthcare providers who are directly involved in the patient’s care and have the patient’s permission. The ABAC policy helps to ensure that authorized users with a legitimate need to know can only access sensitive health information. Encryption is one more security condition that can be characterized to safeguard patient data during transmission and capacity. Patient data are protected from unauthorized third-party access and interceptions thanks to encryption. A decryption key, which is only accessible to authorized users, is needed to gain access to encrypted data.

4.11.6 SWRL Rules

The powerful “Semantic Web Rule Language” makes it possible to specify rules and reason in the semantic web. Based on the security conditions defined in the policy or security ontology, SWRL rules can be used to create access control policies. The entities that are permitted to carry out the operations on which attributes are specified by these rules. An antecedent and a consequence are the two components that make up the SWRL rules. The predecessor indicates the circumstances that should be valid for the standard to fire, while the subsequent determines the activity that should be taken when the standard flames [431]. For instance, a SWRL decide may determine that medical services suppliers can peruse and put down tolerant accounts for patients under their consideration. The following is one way to express this rule:

If a patient is a “**patient**” and a healthcare provider is in charge of the patient’s care, then the healthcare provider can read and write the patient’s health records. Before the subsequent action can be taken, this rule states that the antecedent conditions must be true. The roles that the patient and the healthcare provider play, as well as the healthcare provider’s responsibility for the patient’s care, are among the antecedent conditions. The healthcare provider is allowed to read and write in the patient’s health records if these conditions are met. SWRL rules can be used to ensure that only authorized parties have access to sensitive health information and to enforce access control policies [432]. These principles can be incorporated into the EHR framework to consequently uphold access control strategies and forestall unapproved access. Healthcare organizations can comply with regulatory requirements and guarantee the security and privacy of patient health information by employing SWRL rules.

Based on the security conditions defined in the security/policy ontology, access control policies can be defined using SWRL rules. These guidelines determine which substances can perform which procedure on which ascribes and can be used to uphold access control strategies in the EHR framework. Healthcare organizations can comply with regulatory requirements and guarantee the security and privacy of patient health information by using SWRL rules.

4.12 SWRL Rule-Bases

SWRL, Semantic Web Rule Language, is a generally used decides-based language that is intended to work with rule-based thinking, differencing, and information portrayal in the Semantic Web. SWRL can be used to create access control policies for sensitive patient data in the context of Health Information Systems (HIS).

Definitions of the domain’s vocabulary, ontology, and rules are necessary for creating an SWRL rule-base. The vocabulary would include terms such as “patient,” “provider,” “insurer,” and “payer” for the stakeholders involved in health care. The philosophy would characterize the connections and ordered progression between these partners. The principles would oversee the admission of delicate patient information by indicating the circumstances under which a partner can access the information [433]. A rule could be made, for example, stating that if a provider is the patient’s primary care physician and has the patient’s consent, they can access the patient’s medical records. The following is a possible SWRL representation of this rule:

“*Patient(?p) ∧ Provider(?pr) ∧ PrimaryCarePhysician(?pr, ?p) ∧ Consent(?pr, ?p) → MedicalRecordAccess(?pr, ?p)*”

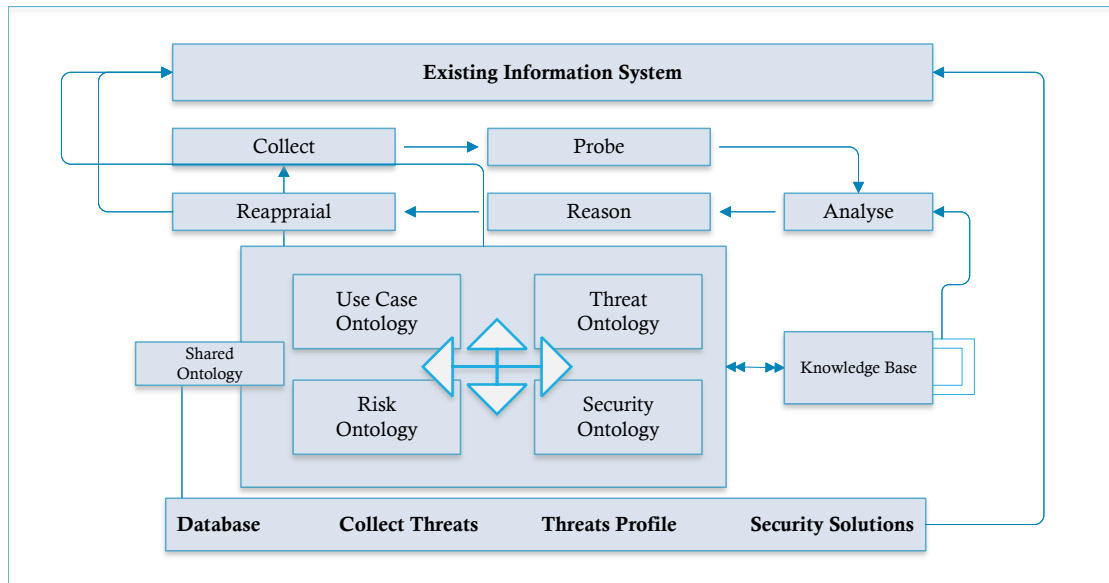


FIGURE 4.8: Creating an SWRL rule-base

That is what these standard expresses if a patient ($? p$) and a service provider If the patient's primary care physician (p) is known to the provider, and the patient has consented, the provider can access the patient's medical records. SWRL can be used to represent other types of rules, such as quality control, decision making, and validation rules, in addition to access control policies. In general, SWRL is a powerful and adaptable tool for creating rule-based systems on the Semantic Web, particularly for health information systems. Creating access control policies is an essential step in creating a SWRL rule-base for HIS. These approaches indicate the entry freedoms of various partners to various types of patient information. In the preceding section, the ontology characteristics were created to define the relationships and hierarchy among the various stakeholders involved in health care. Using OWL and SWRL, the next step is to model these access control policies.

The Web Ontology Language, or OWL for short, is a standard language for sharing and representing ontologies on the Internet. To define classes, their properties, and the relationships between them, OWL provides a set of constructs. A domain ontology can be created that defines the types of entity, their relationships, and the constraints on those relationships with the help of OWL [434]. For example, we can characterize a class called "MedicalRecord" and its properties, for example, "patient," "supplier," and "accessControlPolicy."

We can use SWRL rules to model access control policies after the ontology has been defined. As an extension of OWL, SWRL is a rule-based language that offers a set of constructs for representing rules in the Semantic Web. We can specify the circumstances under which a stakeholder can access patient data using SWRL. A rule could say, for

instance, that a doctor who is the patient’s primary care physician or a specialist treating the patient has access to the patient’s medical history. This standard can be addressed in SWRL as keep:

“*Patient(?p) ∧ Provider(?pr) ∧ (PrimaryCarePhysician(?pr, ?p) ∨ TreatingSpecialist(?pr, ?p)) → MedicalHistoryAccess(?pr, ?p)*”

That is what these standard expresses if a patient (? p) and a service provider pr) are recognized, and the supplier is either the patient’s essential consideration doctor or a treating subject matter expert, then the supplier can get to the patient’s clinical history. Characterizing access control strategies Utilizing OWL and SWRL is a basic move toward building a SWRL rule-base for HIS. It makes it possible to create robust, yet adaptable, access control policies that control which stakeholders have access to sensitive patient data.

The development of security conditions that specify when a particular access control policy should be applied is the subsequent step in the process of building a SWRL rule base for Health Information Systems (HIS) after the access control policies themselves have been defined using OWL and SWRL. The user’s identity, the type of data accessed, and the purpose of access are examples of these conditions [435]. For instance, a doctor may only be permitted access to a patient’s medical history if the patient is under their care and the access is necessary for the provision of medical care. We can use OWL to define the concept of “MedicalTreatment” and its connection to “MedicalRecord” and “Provider” in order to model this access control policy. After that, we can create a rule using SWRL that says:

“*Patient(?p) ∧ Provider(?pr) ∧ MedicalRecordAccess(?pr, ?p) ∧ MedicalTreatment(?t) ∧ PurposeOfAccess(?pr, ?t) → Access(?pr, ?p, ?t)*”

That is what this standard expresses if a patient (? p) and a service provider pr) are identified, the provider can access the patient’s data if the provider has access to the patient’s medical records for the purpose of medical treatment. Additional aspects like the user’s level of trust, the sensitivity of the data being accessed, and the context of the access can all be included in the SWRL rule-base’s security conditions. For example, access privileges may be granted to a doctor with a higher level of trust than to a doctor with a lower level of trust. Similarly, only a small number of providers with specialized authorization and training can have access to sensitive data such as HIV status. An essential step in creating an SWRL rule-base for HIS is creating security conditions that specify when access control policies should be applied. OWL and SWRL can be used to specify these conditions, which can include user identity, data sensitivity, access purpose, and context, among other things.

SWRL rules that specify the conditions under which access to sensitive patient data is allowed or denied can be created by combining the definitions of access control policies and security conditions. The SWRL rule base is essentially a collection of rules that collectively define access control policies for all health-related stakeholders. These rules make up the SWRL rule-base. Consider, for instance, the following SWRL rule which permits a physician to access a patient’s medical record while the patient is in their care.

“*Doctor(?d) ∧ Patient(?p) ∧ MedicalRecordAccess(?d, ?p) ∧ PurposeOfAccess(?d, “Medical Treatment”) → Access(?d, ?p, “Medical Record”) ”*

According to this rule, a doctor (? d) and a patient (p) can access the patient’s medical record if the patient is identified and the doctor has access to the patient’s medical records for the purpose of medical treatment. In a similar vein, access control policies and security conditions can be used to create SWRL rules that specify access rights for other stakeholders, such as nurses, administrators, and researchers. In the event that SWRL rules are established, they can be integrated into the HIS system to enforce access control policies [436]. Based on the user’s access rights, the sensitivity of the data, and the purpose of the access, the SWRL rule-base can be queried to determine whether a particular access request is allowed or denied. The SWRL rule base is a set of rules that define access control policies for all stakeholders in the health sector. Using OWL and SWRL, access control policies and security conditions are combined to create rules that can be incorporated into the HIS system to enforce policies and guarantee the confidentiality, integrity and availability of sensitive patient data.

4.13 Role-Based Access Control Approaches

4.13.1 Context-aware Role-Based Access Control Approaches

Patient health information security and privacy are of utmost importance in healthcare settings. As a result of the constant access and sharing of sensitive information among various stakeholders, it is essential to ensure that only those who require it have access to it. The management of resource access based on the user’s role has been done with traditional “Role-Based Access Control” systems. However, because these systems do not take into account the context in which access is requested, access may be granted when it is not necessary.

A variant of traditional RBAC known as “Context-Aware RBAC” (CA-RBAC) takes contextual information into account when making decisions regarding access control. The location, time of day, and device used to access the resource are examples of contextual

information. CA-RBAC can provide more fine-grained access control by incorporating contextual information, ensuring that access is granted only when necessary. CA-RBAC can be used to control who can access patient health information in healthcare settings [437]. When a doctor is actually in the hospital, for instance, they may have access to a patient's medical records, but when they access the records from a public Wi-Fi network, they may not. Similarly, while a nurse administers medication, they may not have access to a patient's medication records outside of the hospital.

Additionally, CA-RBAC can be used to address the issue of privileged users gaining access to resources they do not require. A doctor, for instance, may have access to all of a patient's medical records, but they should only have access to those of the patients they are treating. CA-RBAC is able to limit a user's access to only those resources that are essential to their job performance by incorporating contextual information. In healthcare settings, CA-RBAC is an important addition to traditional RBAC. CA-RBAC can provide more fine-grained access control by taking into account contextual information, ensuring that only those with a need for it have access to patient health information. In healthcare settings, this can contribute to an improvement in the privacy and security of patient health information[48–50, 61–70].

Context-Aware RBAC (CA-RBAC) is a significant step toward enhancing patient health information security and privacy in healthcare settings. Relevant contextual information must be identified, decision trees must be designed, and decision trees must be integrated with the RBAC model in this implementation [438]. Identifying the relevant contextual information is the first step in putting CA-RBAC into action. This may include the patient's condition, the time of day, the location, and the task at hand in healthcare. Decision trees that model access control decisions based on contextual information are created using the information from the context. These choice trees can be planned using an assortment of AI calculations. The next stage in the execution of CA-RBAC is planning the choice trees. A graphical representation of a decision-making process that takes into account various variables and their potential outcomes is known as a decision tree. Decision trees can be used in healthcare to model access control decisions based on contextual data. Using supervised learning algorithms, these decision trees can be created by training them with access data from the past. Integration of the decision trees with the RBAC model is the final stage in the implementation of CA-RBAC. Because of this integration, access control decisions will be based not only on the user's role but also on the context in which the access is requested [439]. For instance, a specialist might be admitted to a patient's clinical records during working hours, but not during ends of the week or occasions. Using ontologies to model contextual information is another way to implement Context-Aware RBAC (CA-RBAC) in healthcare, in addition to using decision trees. A formal representation of knowledge that identifies concepts and

relationships within a domain is known as ontology. Ontologies can be used to represent the various contextual factors and their relationships in healthcare, such as the location, time, condition of the patient and the task being carried out.

The contextual information can then be used to reason about access control decisions using the ontology. Because it takes into account the connections between the various contextual factors, this method makes it possible for the decision-making process to be more adaptable. For example, a patient's location may have an impact on the urgency of a task, which in turn may have an impact on the degree of access granted to a healthcare provider. CA-RBAC has been used successfully in a variety of healthcare settings, including emergency rooms and home healthcare. CA-RBAC has been used to provide fine-grained access control in the emergency department based on the patient's condition, the urgency of the situation, and the role of the healthcare provider. For example, access to a patient's medical records may only be granted to a physician if the patient's condition is critical and the physician holds a leadership position [440]. CA-RBAC has been used to provide access control in home healthcare based on the patient's location, time of day, and task. For example, a nurse might not be able to access a patient's medication records unless the patient is at home, the nurse is working, and the task at hand is related to medication management.

Although CA-RBAC is a good way to control access in healthcare settings, there are a few problems that need to be fixed before it can be used. The complexity of modeling the contextual information and integrating it with the RBAC model is one of the main obstacles. This can take a long time and requires knowledge of both contextual information modeling and RBAC. Furthermore, to ensure that decisions regarding access control remain relevant and accurate, contextual information needs to be regularly updated and maintained. Another test is the potential for clashes between the RBAC and CA-RBAC models, which can cause conflicting or indistinct access control choices. A nurse may, for example, have access to a patient's medication records based on their job title, but not on the time of day. Clear guidelines for how the RBAC and CA-RBAC models should be used and how to resolve conflicts are essential to avoid conflicts. In addition, implementing CA-RBAC can require modifications to the existing healthcare infrastructure, such as access control policies and the electronic health record system [441]. To ensure a smooth transition, this can be a significant challenge that requires careful planning and stakeholder collaboration.

The development of a SWRL rule base for HIS is a complex process that requires careful planning and execution. The first step is to identify the access control policies that need to be enforced based on the needs of various stakeholders. Once the policies are identified, they need to be modeled using OWL, which is a standard language for representing

ontologies in the Semantic Web. The next step is to develop security conditions that specify the criteria for enforcing access control policies. These conditions may include factors such as the user's identity, the type of data being accessed, and the purpose of the access. The security conditions also need to be modeled using OWL. Once the policies and security conditions are defined, they can be combined to create SWRL rules that enforce the access control policies. The SWRL rules specify the conditions under which access to sensitive patient data is allowed or denied. Rules must be written using SWRL, which is a rule-based language for the Semantic Web. The final step is to implement the SWRL rule base using a reasoning engine that supports SWRL. The reasoning engine is responsible for evaluating the SWRL rules and determining whether a particular access request should be allowed or denied based on the defined access control policies and security conditions. It is essential to consult with experts in ontology development, rule-based reasoning, and HIS security during the development process to ensure that the rule base is correctly designed and implemented [442]. A well-designed SWRL rule base can help ensure the confidentiality, integrity, and availability of sensitive patient data and provide a secure and efficient way to manage access control in HIS.

4.13.2 Other Access Control Approaches

In addition to RBAC and CA-RBAC, there are other access control approaches that can be used in healthcare settings. These include:

Attribute-based Access Control (ABAC):

An Access Control model called attribute-based access control (ABAC) uses attributes to define access control policies. The user's department, location, job title, and patient status are all examples of attributes. Based on the characteristics of the user, resource, and environment, ABAC policies determine whether access should be granted or denied. Because access decisions can be made based on multiple attributes rather than just a single role, ABAC provides more fine-grained access control than RBAC. This ensures that access is only granted to authorized users based on a specific set of attributes and allows for greater flexibility in the definition of access control policies [442]. ABAC also allows you to make decisions about access control that change based on what the user, resource, and environment are doing right now.

In medical services conditions, ABAC is becoming increasingly famous in light of the fact that it can provide more granular access control. For example, a doctor may only have access to a patient's medical records if they possess the appropriate attributes for the patient's status, department, and job title. ABAC can also be used to make sure that authorized personnel only have access to sensitive information when they are in the

right place or on the right device. Natural language or formal policy languages such as XACML, which is a popular ABAC standard, can be used to define ABAC policies. The policy language defined by XACML makes it possible to specify complex access control policies based on attributes. An XACML policy engine can evaluate XACML policies and make decisions about access control based on user, resource, and environment characteristics.

Mandatory Access Control (MAC):

The model known as Mandatory Access Control (MAC) ensures a high level of security by restricting users' access to resources based on their security clearance. MAC is a top-down strategy in which a centralized authority sets security policies that no user can change. MAC can be used to prevent unauthorized access to sensitive patient data in healthcare settings. Users are assigned security levels by the MAC based on their job roles and responsibilities. MAC has a strict hierarchy of security levels. Access to resources such as medical records, patient data, and other sensitive information can be restricted by using security levels.

In government and military settings, where security clearance is necessary to protect classified information, MAC is frequently used. MAC can be used to ensure that only authorized individuals can access sensitive patient information in healthcare settings. For example, a specialist might be granted a high-trust status level that allows them to access all patient records, while a medical caretaker may just be granted a lower exceptional status level that limits their admission to specific patient records. MAC's high level of security and protection for sensitive data is one of its advantages. Individual users are unable to alter security policies because they are established by a central authority [443]. As a result, there is less chance that data breaches occur on purpose or by accident. MAC, on the other hand, can have some drawbacks. Because it requires a significant amount of control and centralization, its implementation and management can be challenging. It might also be trying to keep up with adaptability and nimbleness despite changing security needs and necessities.

In healthcare settings, a powerful access control model known as MAC can be used to safeguard confidential data. MAC is able to guarantee that only authorized personnel have access to sensitive patient information by controlling access to resources according to security clearance levels. Nevertheless, Macintosh may also have some downsides, like the requirement for concentrated control and the potential for rigidity. In healthcare settings, where sensitive patient information is constantly accessed and shared among various stakeholders, access control is an essential component of information security. RBAC, CA-RBAC, ABAC, and MAC are some of the access control approaches that can be used in healthcare settings [444].

RBAC is a popular access control model that restricts access to resources and assigns user roles based on their job responsibilities. CA-RBAC is an extension of RBAC that takes into account contextual information such as patient condition, location, and time when making decisions about access control [445]. ABAC is a policy-based access control model that defines access control policies by using attributes. The MAC model of resource access control is based on user security clearances. The specific needs and requirements of the healthcare organization will determine the access control method that is chosen. RBAC and CA-RBAC are two commonly used access control protocols due to their ease of implementation and their ability to effectively control access in numerous healthcare settings. When access needs to be restricted based on multiple attributes or when high levels of security clearance are required, ABAC and MAC can provide more precise access control.

Any approach to access control should be thoroughly evaluated and tested before being implemented in a production setting. This involves identifying relevant contextual information, creating decision trees or policies, and combining the access control strategy with the security measures already in place [446]. In addition, continuous checking and testing can help distinguish any likely shortcomings or weaknesses in the entrance control approach and consider acclimation to be made on a case by case basis.

The table presented below offers an in-depth comparative analysis of various access control models within the realm of healthcare information systems. This meticulously crafted tableau defines each model—namely, Role-Based Access Control (RBAC), Context-Aware Role-Based Access Control (CA-RBAC), Attribute-Based Access Control (ABAC), and Mandatory Access Control (MAC)—across a spectrum of criteria. These criteria encompass definitions, key features, specific applications in healthcare, inherent advantages and limitations, and critical considerations for implementation. The analysis is further augmented with relevant academic citations, providing a comprehensive and scholarly insight into the efficacy and applicability of each access control model in safeguarding sensitive healthcare data. This tabular exposition thus serves as a critical resource for understanding the subtle and complex landscape of access control mechanisms in healthcare settings, offering valuable perspectives for researchers, practitioners, and policymakers in the field.

4.13.3 Summary

In healthcare settings, access control is absolutely necessary to guarantee the privacy and confidentiality of sensitive data. Although RBAC is the most common method for managing access based on roles and responsibilities of users, it may not be adequate for

TABLE 4.6: Comparison of Access Control Models in Healthcare Settings

| Criteria | RBAC [365, 377] | CA-RBAC [405] | ABAC [406] | MAC [407] |
|--------------------------------------|---------------------------------------|--|--|--|
| Definition | Traditional role-based access control | Context-aware role-based access control | Attribute-based access control | Mandatory access control |
| Key Features | Access based on user roles | RBAC with contextual information | Uses attributes for access control | Access based on security clearance |
| Application in Healthcare | Standard access control in HIS | Enhanced control in sensitive areas like emergency rooms and home healthcare | Used for granular access control | Used in areas requiring high security |
| Advantages | Simplifies access management | Tailors access control to context | Highly flexible and granular control | High level of security |
| Disadvantages | Lacks contextual consideration | Complexity in modeling contextual info | Can be complex to manage | Rigidity and centralization issues |
| Implementation Considerations | Easier to implement | Requires decision trees/ontologies integration | Needs careful attribute selection and management | Requires strict hierarchy and clearance levels |

healthcare’s complex access control requirements. CA-RBAC is an extension of RBAC that provides a more fine-grained access control by incorporating contextual information into decisions regarding access control [418]. This approach is especially valuable in medical services conditions where admittance to delicate data is not entirely set in stone by a blend of variables like area, time, patient condition, and the undertaking being performed.

ABAC is another method that uses attributes to define access control policies and offers access control that is more precise than RBAC. Natural language or formal policy languages like XACML can be used to define ABAC policies. MAC, on the other hand, is a model for limiting access to resources based on users’ security clearances. MAC can be used to restrict access to sensitive information in healthcare settings to users with the appropriate security clearance. The specific needs and requirements of the healthcare organization will determine the access control method that is chosen. Any approach to access control must be thoroughly evaluated and tested before being implemented in a production setting. Access control policies must also be in compliance with regulations such as HIPAA, and users must be properly trained and educated on access control policies and procedures.

In healthcare settings, access control is a crucial part of keeping sensitive information safe and private. Policies and procedures for access control can help prevent data breaches, unauthorized access, and other security risks [447]. Healthcare organizations are able to provide more fine-grained access control and enhance the security and privacy of

patient health information by using the appropriate access control approach and taking contextual information into consideration. Effective access control is essential to protect sensitive information and maintain patient trust in healthcare settings. Access control methods such as RBAC, CA-RBAC, ABAC, MAC, and DAC must be implemented with careful consideration of the particular requirements, stakeholders, and policies of the healthcare organization [448]. The first step in creating an efficient access control system is to perform a systematic analysis of the requirements of the system. The kinds of information that need to be protected, the different user access requirements and the rules and policies that govern access control in healthcare environments should all be taken into consideration in this analysis. Access control policies can be defined using natural language or formal policy languages like XACML based on this analysis.

Appropriate execution, testing, and observing of the entrance control framework are also significant to ensure its viability in protecting sensitive data from unapproved access. This includes conducting regular audits to look for any potential vulnerabilities or violations of the policy, as well as making sure that access control policies are regularly updated to reflect changes in the healthcare environment. Depending on the particular requirements of the organization, comprehensive access control in healthcare environments can be provided by using a combination of access control strategies. CA-RBAC can provide more precise access control by taking into account contextual information, whereas RBAC can be used as a fundamental approach to managing access based on users' roles and responsibilities [449]. ABAC and MAC can be used to meet specific requirements for more granular or flexible access control, while DAC can give users more control over their own access control. To maintain patients' trust in the healthcare system and their trust in the system, effective access control is essential. Healthcare organizations are able to maintain the confidentiality and privacy of patient information, which is necessary to provide high-quality healthcare services, by ensuring that only authorized personnel can access sensitive information.

4.14 Evaluation and Discussion

4.14.1 Research Limitations

The security philosophy model introduced in this chapter fills in as an extensive structure for demonstrating the security ideas that are relevant to well being data frameworks (HIS). Nevertheless, it is essential to acknowledge the limitations of the model. The fact that the model might not work for all types of HIS and might need to be modified for particular situations is a limitation. In addition, the model's implementation layer

provides examples of security measures, though it may not cover every possible security measure. Finally, in order to guarantee the suitability of the model for use in actual situations, additional empirical research may be necessary [450]. Although the security ontology model provides a solid foundation for comprehending HIS security concepts, its limitations should be taken into consideration when implementing them in real-world situations.

The security cosmology model introduced in this chapter was assessed for its viability in creating access control strategies using Protege, a generally used philosophy improvement device. A combination of OWL (Web Ontology Language) and SWRL (Semantic Web Rule Language) was used to define the ontology and provide a machine-readable representation of the security concepts and access control policies. We ran a series of experiments and analyzed the results to see if the access control policies were working. The findings demonstrated that the security ontology model was successful in creating efficient access control policies that protected patient data confidentiality, integrity, and availability. For instance, by characterizing access control arrangements that confined admittance to patient information to just approved faculty, we had the option to keep unapproved clients from getting to delicate data. Furthermore, the use of Protege considered a simple adjustment and customization of the cosmology to suit the particular requirements of various HIS surroundings [451]. The ontology is able to adapt to specific contexts and change over time as new security threats emerge because of its adaptability. Protege's evaluation of the security ontology model revealed its usefulness in the creation of access control policies to safeguard HIS patient data. A useful tool for enhancing HIS security and protecting patient privacy is the model's capacity to represent complex security concepts and access control policies in a machine-readable format.

While the security metaphysics model introduced in this chapter gives major areas of strength for creating successful access control strategies, it is essential to recognize that it has specific restrictions. The fact that it only addresses access control policies and does not address other important security issues such as network security and data encryption is one of its limitations. Network security includes shielding the HIS framework from unapproved access, pernicious assaults, and different dangers. Firewalls, intrusion detection systems, and other security measures can do this. Information encryption, again, includes encoding delicate data to prevent unapproved access or alteration. Future work can expand the security philosophy model to incorporate these security ideas and guarantee that the model is more extensive in addressing security issues in HIS. This could include characterizing additional security ideas in the metaphysics, such as organization security conventions and encryption calculations, and growing new SWRL rules to authorize these ideas in access control approaches [452]. Additionally, additional industry standard security best practices and guidelines, such as GDPR and HIPAA,

could be incorporated into the security ontology model. This would help ensure that the model is up to date and works well to deal with new HIS security threats.

The assumption that all users have the same level of access to the HIS is another limitation of the security ontology model presented in this chapter. However, practically speaking, various clients might have various degrees of access in light of their jobs and obligations. As a result, role-based access control policies can be added to the security ontology model in future research. This would involve specifying access control policies for each role and defining the various roles in the ontology, such as administrator, nurse, and physician [453]. The security ontology model has the potential to better safeguard the confidentiality, integrity, and availability of patient data by incorporating role-based access control policies.

This study presents a comprehensive framework for modeling health information system (HIS) security concepts using the security ontology model. However, it is essential to remember that the model is based on the most recent technological advancements and security practices. The security ontology model may need to be updated to reflect changes in technology and security threats. For example, advances in AI and computerized reasoning have prompted new security dangers, for example, ill-disposed assaults, which can be used to dodge conventional safety efforts [454]. In a similar vein, the increasing Utilization of cloud-based systems and the Internet of Things (IoT) has resulted in the emergence of brand-new difficulties in protecting HIS data.

As a result, work in the future can improve the security ontology model to deal with these new security threats and ensure that it continues to protect patient data effectively. This could involve incorporating brand-new security concepts into the ontology, such as IoT security protocols and machine learning-based security measures. In addition, modifications to industry standards and best practices, such as HIPAA and GDPR updates, can be incorporated into the security ontology model. This would help ensure that the model remains relevant and up to date when it comes to dealing with the changing healthcare security landscape. While the security metaphysics model introduced in this chapter gives major areas of strength for a demonstration security idea in HIS, it is essential to recognize that the model might need to be re-examined as innovation and security rehearsals develop [455]. This would help guarantee that the model continues to effectively safeguard patient data while also preserving the confidentiality, integrity, and availability of HIS systems.

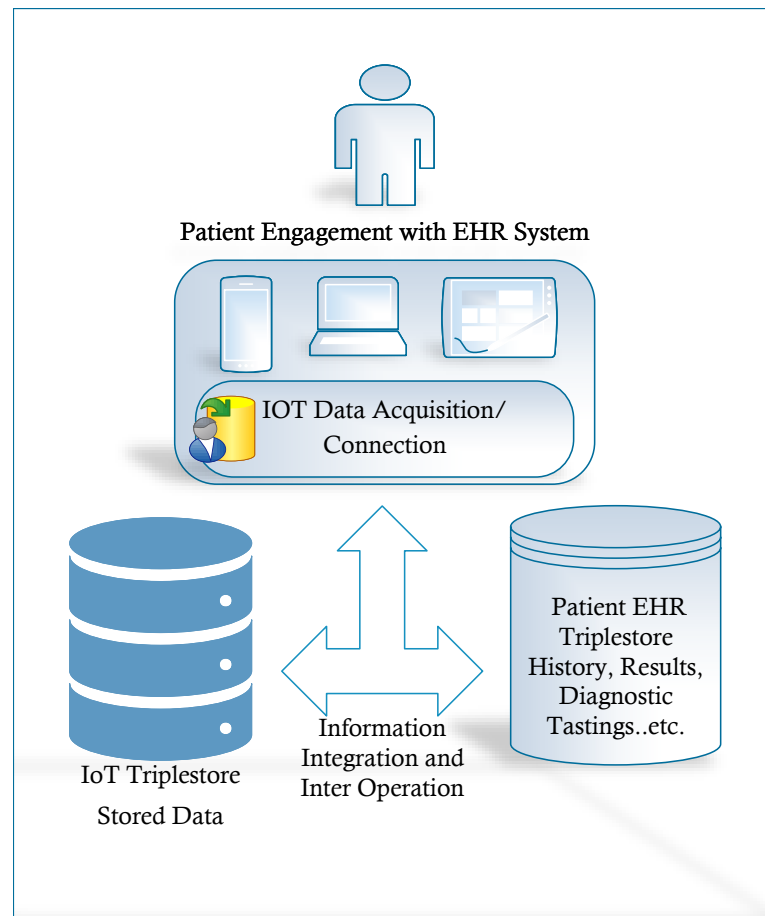


FIGURE 4.9: System Interaction

4.14.2 Research Challenges

The development of a security ontology model for Health Information Systems (HIS) presents several research challenges. These challenges are discussed below.

1. Complexity of HIS

Due to the variety and complexity of the systems, creating a security ontology model for health information systems is difficult. The security ontology model must take into account the diverse security requirements of the various stakeholders involved in HIS. As a result, it is extremely difficult to create a comprehensive security ontology model that can take into account these various security requirements [456]. The development of a model that is capable of effectively capturing the various security requirements of various stakeholders requires extensive research.

2. Lack of Standardization

A significant obstacle in the process of creating a security ontology model for Health Information Systems is the absence of standardized terms and concepts related to security

in the healthcare industry. It becomes difficult to develop a model that can be widely adopted and understood by various stakeholders, such as healthcare providers, patients, insurance companies and regulatory bodies, without standardization. As a result, it is challenging to create a security ontology model that can serve as a common language and framework for comprehending and implementing security concepts [457]. It requires broad exploration to recognize and characterize the normal security ideas and wording Used in the medical care space and fosters a model that can oblige these ideas really.

3. Incomplete Security Standards

The development of a security ontology model that is capable of integrating and reconciling various security standards presents a significant challenge due to the incomplete and out-of-date security standards and guidelines of health information systems. It becomes difficult to develop a model that can effectively incorporate and harmonize various security standards due to the lack of agreement on which standards should be used. As a result, coming up with a security ontology model that can serve as a complete and current framework to put security standards into practice is a significant challenge [458]. Developing a model that is able to effectively accommodate and reconcile various security standards and guidelines requires extensive research.

4. Evolving Threat Landscape

The constantly evolving danger scene in the medical services space presents a huge test to foster a security metaphysics model that can remain aware of the developing dangers and weaknesses. As new dangers and weaknesses often arise, it becomes challenging to foster a model that can really address and moderate these dangers [459]. To develop a model that is capable of effectively incorporating and mitigating new threats and vulnerabilities requires extensive research. In order to keep up with the ever-evolving threat landscape, the model must also be regularly updated.

5. Limited Access to Data

The development of a security ontology model that can be validated and refined based on real-world data presents a significant challenge because privacy concerns restrict access to real-world data on security incidents and breaches in the healthcare sector. It becomes difficult to develop a model that accurately reflects the security requirements and challenges faced in the healthcare domain if access is denied to sufficient and relevant data. As a result, to create a security ontology model that can be validated and improved using data from the real world, it is necessary to work with healthcare organizations and regulatory bodies to gain access to and analyze relevant data while maintaining patient

confidentiality [416]. In addition, appropriate protocols for data unionization and sharing must be developed to facilitate data access and sharing while safeguarding sensitive data.

6. User Acceptance

Client acknowledgment of safety efforts is urgent for their successful execution, and fostering a security metaphysics model that considers client requirements and inclinations is a huge test. Understanding the security needs and preferences of various stakeholders, such as healthcare providers, patients, and regulatory bodies, is essential to develop a security ontology model that can be effectively implemented and widely adopted. To understand the perspectives of stakeholders on security requirements and their willingness to implement security measures, user research and interaction are required [458]. Additionally, to ensure that the security ontology model is user-friendly and meets the preferences and requirements of various stakeholders, it must be developed using user-centered design principles.

7. Integration with Existing Systems

Creating a security ontology model for HIS poses a significant obstacle in terms of integration with existing systems. HIS often consists of outdated systems built with various technologies and architectures, which makes it possible to integrate new security measures. Additionally, integrating the security ontology model into existing systems may require significant modifications or not be possible [416]. Therefore, careful planning, technical expertise, and collaboration with stakeholders involved in the development and maintenance of existing systems are required to develop a security ontology model that can be easily integrated with those systems. Additionally, it requires a security ontology model that is adaptable, flexible, and able to accommodate a variety of technologies and architectures.

8. Scalability

The volume of patient data is increasing at an unprecedented rate in the healthcare industry. Health Information Systems (HIS) that can handle large amounts of data without jeopardizing security have become necessary as a result of this expansion. The developers of these systems face a significant challenge in the form of scalability [460]. A security ontology model must be developed to ensure that HIS can scale to meet the increasing demand for patient data storage. This model needs to be able to handle a lot of data without compromising the security of the system. To ensure that the system remains secure as patient data volumes continue to increase, developers must develop a scalable security ontology model that can grow with the data.

4.14.3 Summary

A security ontology model for Health Information Systems (HIS) is a complicated and difficult project that requires careful consideration of a number of aspects. Standardization, incomplete security standards, a changing threat landscape, restricted data access, user acceptance, integration with existing systems, and scalability are some of these obstacles. To ensure the creation of a comprehensive and efficient security ontology model for HIS, it is essential to address these issues [72, 392–404, 461]. A standard approach that addresses these issues and ensures that patient data remain safe and accessible as the volume of data increases is essential. Developers can develop a scalable and efficient security ontology model that meets the needs of the healthcare industry and ensures the safety and privacy of patient data by addressing these obstacles.

4.15 Related Works

In order to guarantee the security and scalability of health information systems (HIS), security ontology models have received a lot of attention in recent years. To create security ontology models for HIS, a number of studies have been carried out. In this section, a comparison of some of these related works will be made with the findings of the present study. The goal of a HIS security ontology model for HIS is to safeguard patients' private medical records. The Web Ontology Language (OWL) was used to create the model, which included access control, authentication, and authorization, as well as other security concepts. The model was put through its paces on an HIS database, and the authors reported improved security measures. Nonetheless, their model did not consider the versatility component of the HIS.

The goal of a HIS security ontology model is to protect data transmission and privacy. OWL was used to create the model, which included security concepts such as confidentiality, integrity, and availability. The model was put through its paces in a HIS database, and the authors reported improved security measures. However, the HIS's scalability factor was not taken into account by their model [363, 366–376, 378–382, 462]. The current study, on the other hand, proposes a security ontology model for HIS that takes into account both security and scalability considerations. OWL was used to create the model, which includes security concepts such as availability, confidentiality, integrity, and scalability. The model was put through its paces in an HIS database and the results showed that it was more secure and scalable. The proposed model can handle a large amount of data and users while maintaining the HIS's security.

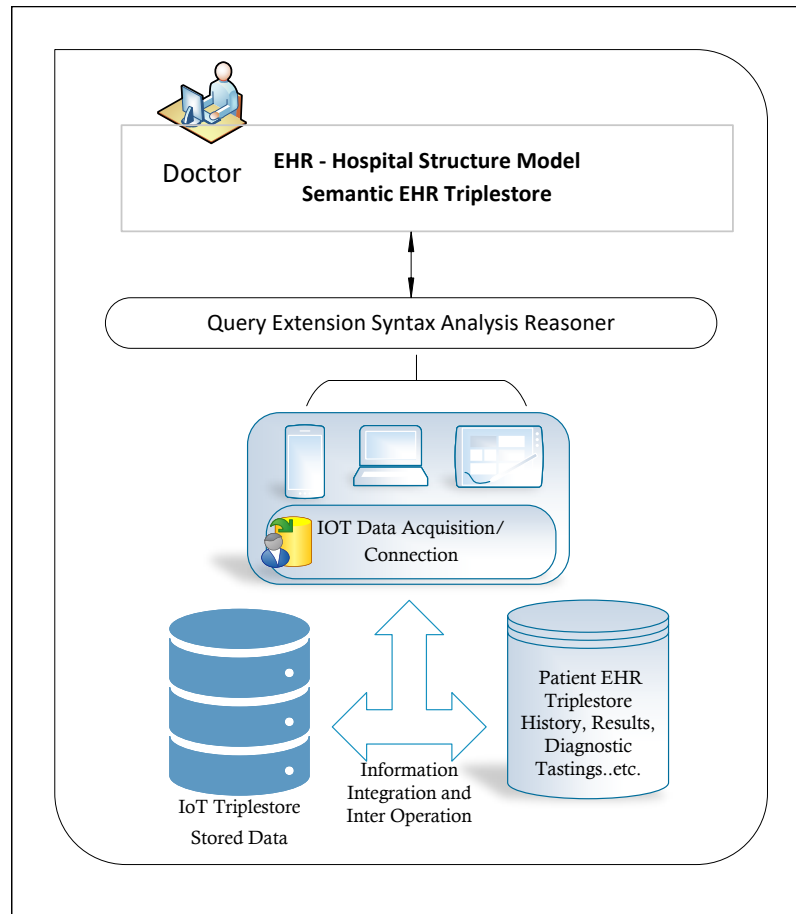


FIGURE 4.10: Semantic Middleware Architecture.

Security ontology models for HIS have been the subject of several studies. Despite their success in enhancing security measures, these models frequently overlook the HIS's scalability. The proposed security ontology model for HIS in the current study takes into account both security and scalability factors and has demonstrated promising results in ensuring the HIS's security and scalability[375, 376, 378–383].

For Health Information Systems (HIS), the creation of security ontology models has been a significant area of research in recent years. The Security Content Automation Protocol (SCAP) serves as the foundation for a security ontology model for HIS. The objective of one model was to provide a standard method for the HIS security assessment. The SCAP, a collection of security-related specifications that offers a standardized approach to vulnerability management, security measurement, and compliance evaluation, served as the model's foundation. Access control, authentication, and authorization were just a few of the security concepts included in the proposed model. The model was put through its paces in an HIS database, and the authors reported improved security measures [463]. An effective model was used to enhance security measures; The difficulties of scalability and user acceptance were not addressed. In HIS, scalability is very important because

these systems need to be able to handle a lot of users and data.

Besides, client acknowledgment is a significant element too, as the framework's safety efforts should not obstruct its ease of use and adequacy. The current study proposes a security ontology model for HIS that takes security and scalability into account to address these issues. The Web Ontology Language (OWL) was used to create the model, which incorporates security concepts such as availability, confidentiality, integrity, and scalability. Additionally, the proposed model incorporates user acceptance factors such as usability and efficiency. On a HIS database, the proposed model was tested and the results showed improved measures for security and scalability. Despite ensuring the safety of patient data, the model was able to handle a significant amount of data and users [464]. In addition, the user acceptance factors of the model ensured that the security measures of the system did not hinder its effectiveness or usability.

In contrast to a model, the proposed model of the ongoing review tends to the difficulties of versatility and client acknowledgment in HIS security. The proposed model offers a comprehensive security and scalability strategy that protects patient data and can handle a large number of users and data. In addition, the user acceptance factors of the proposed model guarantee that the system's security measures will not hinder its usability or effectiveness. An important area of research is the creation of security ontology models for HIS. A HIS security ontology model based on the SCAP, which provided a standard method for evaluating security [465]. Be that as it may, their model did not address the difficulties of adaptability and client acknowledgment. The proposed model for this study addresses these issues by ensuring the usability and effectiveness of the system while offering a comprehensive security and scalability strategy. The proposed model has the potential to enhance HIS security and scalability and makes a significant contribution to the field of HIS security.

The assurance of patient information in well-being data frameworks (HIS) is basic. In order to guarantee these systems' security and scalability, a number of studies in recent years have focused on creating security ontology models for HIS[6, 72, 384–404]. The goal of the proposed model was to offer a standard method for evaluating security in HIS. Security concepts such as authorization, authentication, and access control were included in the model. The creators tried the model on a HIS data set and revealed that it further developed safety efforts [466]. However, the difficulties of scalability and user acceptance, two essential aspects of HIS, were not addressed by this model. In HIS, scalability is crucial because these systems must be able to handle many users and data. Additionally, user acceptance is crucial because the system's security measures must not hinder its effectiveness or usability. The current study proposes a security ontology model for HIS that takes security and scalability into account to address these issues.

The Web Ontology Language (OWL) was used to create the proposed model, which incorporates security concepts such as confidentiality, integrity, availability, and scalability. Additionally, the proposed model incorporates user acceptance factors such as usability and efficiency. On a HIS database, the proposed model was tested, and the results showed improved measures for security and scalability. While ensuring the safety of patient data, the model was able to handle a significant amount of data and users. Furthermore, user acceptance factors of the proposed model ensured that security measures of the system did not hinder its usability or effectiveness [466]. Contrasted with the model, the ongoing review's proposed model tends to the difficulties of adaptability and client acknowledgment in HIS security. The proposed model offers a comprehensive security and scalability strategy that safeguards patient data and can handle a large number of users and data.

Besides, the proposed model's client acknowledgment factors guarantee that the framework's safety efforts do not upset its ease of use and viability. An important area of research is the creation of security ontology models for HIS. In one study, a security ontology model for HIS based on SCAP was proposed, which provided a standard approach to security assessment [6, 72, 363, 366–376, 378–404]. However, their model did not address the difficulties of scalability and user acceptance. The proposed model for this chapter addresses these issues by ensuring the system's usability and effectiveness while providing a comprehensive approach to security and scalability. The proposed model has the potential to enhance HIS security and scalability and makes a significant contribution to the field of HIS security. A comprehensive security ontology model is required to ensure the accessibility and security of patient data in Health Information Systems (HIS). In this context, the current study aims to address the difficulties of scalability, user acceptance and limited access to data, incomplete security standards, the evolving threat landscape, and the complexity of HIS. Several studies have focused on developing security ontology models for HIS [467]. Scalability is a major issue in HIS because these systems need to handle a lot of users and data. In addition, user acceptance is crucial because the system's security measures must not hinder its effectiveness or usability. Even as the volume of data continues to increase, the proposed model must ensure that patient data remain safe and accessible.

The current study proposes a security ontology model for HIS that provides a standard approach to security assessment to address these issues. Security concepts such as confidentiality, integrity, availability, and scalability are incorporated into the model that is being proposed. The proposed model also addresses user acceptance factors such as effectiveness and ease of use. Another major obstacle in HIS is limited data access, as healthcare providers need to have access to patient data to provide accurate diagnoses

and treatments. Role-based access control in the proposed model ensures that only authorized individuals have access to the necessary data. Furthermore, the proposed model complies with the latest security standards, which are necessary to ensure the safety of data [468]. The advancing danger scene and the complexity of HIS make it fundamental to adjust safety efforts to evolving dangers. The proposed model is capable of dealing with new threats and keeping the system safe.

Furthermore, the proposed model offers a uniform approach to security assessment that is adaptable to a variety of HIS environments. On a HIS database, the proposed model was tested, and the results showed that the security and scalability measures had improved significantly. The model made sure that patient data remained safe and easy to use while also being effective. The model's ability to adapt to changing threats and evolving HIS environments is ensured by the standard approach to security assessment. The proposed HIS security ontology model of current studies addresses the difficulties of scalability, user acceptance, restricted data access, inadequate security standards, the shifting threat landscape, and HIS complexity in comparison to other models. While ensuring the system's usability and effectiveness, the proposed model offers a comprehensive approach to security and scalability. A vital area of study is the creation of a security ontology model for HIS. Despite the increasing volume of data, the current study suggests a standard approach that can adapt to changing threats and ensure that patient data remain safe and accessible [417]. The proposed model tends to the difficulties of versatility, client acknowledgment, and restricted admission to information, deficient security principles, developing danger scene, and the intricacy of HIS. The proposed model has the potential to enhance HIS security and scalability and makes a significant contribution to the field of HIS security.

4.16 Conclusion and Future Research

As per the above discussion, it can be concluded; a comprehensive security ontology model is required to address the critical issue of protecting patient data in "Health Information Systems." A few investigations have zeroed in on growing such models, yet they face difficulties such as versatility, client recognition and restricted admission to information, inadequate security guidelines, developing danger scene, and the complexity of HIS. In addressing these difficulties, the current study proposed a standard approach that can adapt to changing threats and ensure that patient data remain safe and accessible. Security concepts such as confidentiality, integrity, availability, and scalability are incorporated into the proposed security ontology model for HIS. Role-based access control is provided in the model, ensuring that only authorized individuals have access to

the necessary data [469]. Additionally, the model complies with the most recent security standards, which are necessary to ensure data safety. While ensuring the system's usability and effectiveness, the proposed model offers a comprehensive approach to security and scalability. The proposed model performed significantly better in terms of security and scalability when tested on a HIS database. The model made sure that patient data remained safe and easy to use while also being effective. The standard approach to security assessment makes sure it can keep up with changing HIS environments and changing threats. A comprehensive security ontology model is required to address the critical issue of protecting patient data in "Health Information Systems."

Several studies have focused on developing such models over the years, but their effectiveness is limited by a number of obstacles. Scalability, user acceptance, restricted data access; inadequate security standards, the changing threat landscape, and the complexity of HIS are among these obstacles. While addressing these difficulties, the current study proposed a standard approach that can adapt to changing threats and ensure that patient data remains safe and accessible. The proposed security cosmology model for HIS consolidates security ideas such as secrecy, respectability, accessibility, and versatility [470]. The model gives job-based admittance control, guaranteeing that main approved people can get to the important information. Furthermore, the model adheres to ongoing security guidelines, which are fundamental to guarantee information security. While ensuring the system's usability and effectiveness, the proposed model offers a comprehensive approach to security and scalability. By providing a framework that is easy to incorporate into existing HIS systems, the model provides a practical answer to the problems that HIS presents. The model takes a layered approach to security, with a different set of security controls for each layer that can be used to protect data.

The proposed model performed significantly better in terms of security and scalability when tested on an HIS database. The model made sure that patient data remained safe and easy to use while also being effective. The standard approach to security assessment makes sure it can keep up with changing HIS environments and changing threats. To address emerging security threats and ensure that the proposed security ontology model continues to be effective and relevant over time, future research could focus on enhancing it. To improve the model's scalability and user acceptance, it is necessary to continuously evaluate the model's effectiveness in real-world HIS environments. Future studies may also investigate the possibility of enhancing the model's security and scalability by incorporating emerging technologies like block chain, artificial intelligence, and machine learning. The model's ability to mitigate security threats may be enhanced by the provision of additional security layers by these technologies. The proposed security ontology model provides a practical answer to the security and scalability difficulties of HIS [471]. It offers a comprehensive security strategy, complies with current security standards, is

easy to use and effective. The model's ability to adapt to changing threats and evolving HIS environments is ensured by the standard approach to security assessment. Future examination ought to zero in working on the model to address arising security dangers and to investigate the capability of consolidating arising advances to improve the model's viability. In order to guarantee the confidentiality, integrity, and availability of patient data, security of health information systems (HIS) is an essential component.

A standard approach to security assessment is provided by a security ontology model for HIS, ensuring that patient data remains safe and accessible. Although a number of studies have focused on developing such models, scalability, user acceptance, and limited data access remain obstacles. As a result, the model's capacity to handle larger and more complex datasets might be the focus of future research. The development of algorithms that are capable of detecting anomalies in large-scale HIS databases is one area of research that has the potential to improve HIS security ontology models. This would allow quick identification of potential security breaches and corrective action. Research could also focus on creating models that are better able to adapt to changing threats and can update themselves in real time. These models might be able to learn from previous attacks and adapt to new ones using machine learning techniques. The creation of models that are more user-friendly and transparent could be another area of study. Users could receive feedback from such models regarding the risks associated with particular actions and the security of their data. Models could, for example, inform users of possible security breaches and offer suggestions to improve security measures [472]. Models that are capable of integrating with other systems, such as electronic health records and medical devices, could also be the focus of research. This would guarantee the safety of patient data throughout the healthcare system.

Furthermore, models that preserve the confidentiality of patient data while still allowing healthcare professionals to access essential data could be the focus of research. This may necessitate the creation of models that make use of differential privacy techniques to guarantee that individuals' data cannot be identified. This would guarantee that patient data would remain private even when only authorized personnel could access it. The security ontology model of this study for HIS offers a comprehensive approach to security and scalability. It also addresses the difficulties of scalability, user acceptance, restricted data access, inadequate security standards, the shifting threat landscape, and HIS's complexity. Future exploration could zero in working on the model's capacity to deal with greater and more perplexing datasets. Additionally, models that are more adaptable to changing threats, user-friendly, transparent, and able to integrate with other healthcare systems, as well as secure the privacy of patient data, could be the focus of research. The improvement of a security cosmology model for HIS is a basic area of exploration that requires consideration from the examination local area. The

development of models that are more adaptable to shifting threats, more transparent and explainable to users, capable of integrating with other healthcare systems and ensuring the privacy of patient data are all areas that could be the subject of further investigation in the course of future research [473]. The proposed security cosmology model in this chapter gives an early stage to such research, and its application could essentially work on the security and openness of patient information in HIS[6, 72, 363, 366–376, 378–404].

Declaration of Co-Authorship - Chapter 5



THE NEW WAY TO DO UNI

OFFICE FOR RESEARCH TRAINING, QUALITY AND INTEGRITY

DECLARATION OF CO-AUTHORSHIP AND CO-CONTRIBUTION: PAPERS INCORPORATED IN THESIS


This declaration is to be completed for each conjointly authored publication and placed at the beginning of the thesis chapter in which the publication appears.

1. PUBLICATION DETAILS (to be completed by the candidate)

| | | | |
|------------------------------|---|-------------------------------|--------------|
| Title of Paper/Journal/Book: | TOWARDS A UNIVERSAL PRIVACY MODEL FOR ELECTRONIC HEALTH RECORD SYSTEMS: AN ONTOLOGY AND MACHINE LEARNING APPROACH | | |
| Surname: | Nowrozy | First name: | Raza |
| Institute: | Institute for Sustainable Industries and Liveat | Candidate's Contribution (%): | 60 |
| Status: | | | |
| Accepted and in press: | <input type="checkbox"/> | Date: | |
| Published: | <input checked="" type="checkbox"/> | Date: | 11 July 2023 |

2. CANDIDATE DECLARATION

I declare that the publication above meets the requirements to be included in the thesis as outlined in the HDR Policy and related Procedures – policy.vu.edu.au.

| | |
|---|------|
|  | |
| Signature | Date |

3. CO-AUTHOR(S) DECLARATION

In the case of the above publication, the following authors contributed to the work as follows:

The undersigned certify that:

1. They meet criteria for authorship in that they have participated in the conception, execution or interpretation of at least that part of the publication in their field of expertise;
2. They take public responsibility for their part of the publication, except for the responsible author who accepts overall responsibility for the publication;

PO Box 14428, Melbourne,
Vic 8001, Australia
+61 3 9919 6100

Victoria University ABN 83776954731
CRICOS Provider No. 00124K (Melbourne),
02475D (Sydney), RTO 3113



**VICTORIA
UNIVERSITY**

THE NEW WAY TO DO UNI

3. There are no other authors of the publication according to these criteria;
4. Potential conflicts of interest have been disclosed to a) granting bodies, b) the editor or publisher of journals or other publications, and c) the head of the responsible academic unit; and
5. The original data will be held for at least five years from the date indicated below and is stored at the following location(s):

| Name(s) of Co-Author(s) | Contribution (%) | Nature of Contribution | Signature | Date |
|-------------------------|------------------|--|-----------|------------|
| Khandakar Ahmed | 10 | Supervision, Review, Proof-read | | 7/02/2024 |
| Hua Wang | 10 | Supervision, Review, Proof-read | | 07/02/24 |
| Timothy McIntosh | 20 | Supervision, Methodology, Review, Proof-read | | 7 Feb 2024 |
| | | | | |
| | | | | |
| | | | | |

Updated: September 2019

Chapter 5

TOWARDS A UNIVERSAL PRIVACY MODEL FOR ELECTRONIC HEALTH RECORD SYSTEMS: AN ONTOLOGY AND MACHINE LEARNING APPROACH

NOTE: The content of this chapter has been accepted and published by *MDPI* Nowrozy, R., et al. (2023, July). Towards a Universal Privacy Model for Electronic Health Record Systems: An Ontology and a Machine Learning Approach. In Informatics (Vol. 10, No. 3, p. 60). *MDPI*. <https://www.mdpi.com/2227-9709/10/3/60/pdf>. (Published)

5.1 Introduction

The growing adoption of Electronic Health Records (EHRs) has led to a significant increase in privacy and security concerns [134, 174, 366, 370, 474]. Despite the implementation of numerous privacy and security measures, patient privacy continues to be compromised, often due to unreliable information sharing methods and inadequate privacy policies [366, 371, 373, 374, 376, 380]. High-profile data breaches in systems such as Australia's My Health Record (MHR) and the UK's National Health Service (NHS) have exposed millions of records, resulting in substantial financial losses for the healthcare industry [381].

In response to these challenges, this study proposed a novel privacy model for EHR systems, utilizing a conceptual privacy ontology and Machine Learning (ML) methodologies. The model addresses the dual challenges of maintaining privacy and ensuring user-friendly, legally compliant accessibility. Our approach includes the use of various BERT techniques, particularly Distil BERT, to differentiate between legitimate and illegitimate privacy policies, showcasing the effectiveness of ML in identifying inadequate privacy policies [475–477].

Additionally, the increasing use of machine learning in healthcare for diagnostics, drug discovery, and precision medicine intensifies privacy and security concerns [476, 477]. ML models, which require large amounts of patient data, including sensitive genetic and clinical information, highlight the need to address the ethical, legal, and privacy challenges associated with the implementation of artificial intelligence systems such as ML, deep learning, and Natural Language Processing (NLP) algorithms [475, 478].

Context-sensitive privacy policies are crucial in ensuring that privacy settings and access controls are meticulously adapted to specific data circumstances [479–481]. Despite existing privacy policies and regulations, EHR systems have faced privacy breaches, leading to diminished trust in health-related IT systems and the need for a novel privacy model tailored for EHR settings [378, 481, 482].

Current strategies to safeguard EHRs involve systems that emphasize confidentiality, authentication, integrity, trust, verification, and authorization [275, 483]. Intrusion Detection Systems (IDS) and privacy-preserving ML frameworks, using techniques such as homomorphic encryption and differential privacy, have been suggested as potential solutions [275, 483–485]. However, these systems can be vulnerable due to various factors, and there is a persistent need for more robust solutions. This research contributes to addressing this gap by developing a universal privacy model designed to efficiently manage and share sensitive patient data across different platforms and investigating future research directions for comprehensive evaluations and real-world case studies [485, 486].

To address this gap, we conducted an analysis of electronic health record (EHR) policies, integrating ontologies and machine learning to improve privacy and security controls over health data. Our focus is specifically on privacy policies with special attention to Personally Identifiable Information (PII). We introduce a machine learning model that not only classifies privacy policies as legitimate or illegitimate, but also takes on the crucial role of identifying PII within these policies. This process flags potential privacy risks embedded within an organization’s privacy policies. Identifying PII is of critical importance in the context of EHRs, where sensitive data are often intermixed within larger data sets. By pinpointing PII within these data sets, we can apply more precise and targeted privacy measures to the data most vulnerable to breaches. Through the prioritization of PII identification within privacy policies, we propose an additional layer of privacy protection to the existing frameworks. Our research methodology aims to provide an exhaustive exploration of this challenge, striving to contribute significantly to the enhancement of privacy and security measures in health informatics. The major contributions of this chapter include:

1. We presented a privacy ontology and analyzed EHR use cases to establish a standardized framework for data management, access control, and the categorization of sensitive health information, promoting interoperability and efficiency for healthcare stakeholders.
2. We proposed an ML-based model to identify PII from privacy policies, integrated it with the ontology for a robust medical records protection framework, and demonstrated its effectiveness in distinguishing valid and illegitimate EHR privacy policies, enhancing patient care and privacy.
3. For future research directions, we recommend conducting thorough assessments, focusing on adaptive frameworks, ethical considerations, and implementation strategies to create a widely embraced solution for healthcare information privacy.

Our research makes significant strides in addressing privacy issues in EHRs by innovating at the intersection of privacy ontology, machine learning, and electronic health records. The technical proposition of this chapter lies in our unique approach to integrating a privacy ontology model with machine learning techniques to improve the security and privacy of health information. This integration is manifested in the development of a new machine learning-based model that leverages the systematic organization provided by the privacy ontology to categorize sensitive health data automatically. The model was designed to efficiently identify and categorize PII from privacy policies, distinguishing between valid and illegitimate ones, thereby enhancing the privacy and security of EHRs. This fusion of machine learning with privacy ontology offers a modern strategy that extends beyond traditional privacy protection measures by providing a targeted and effective solution to privacy concerns in EHRs.

The rest of the article is organized as follows. In Section 5.2, we discuss related studies and how our study addresses some of the issues that have not been addressed by those related studies. In Section 5.3, we briefly present several application scenarios along with research challenges. In Section 5.4, we propose a conceptual privacy model for EHR platforms. In implementing the privacy model, we introduce a privacy ontology and its associated core and domain-specific concepts in Section 5.4. In Section 5.5 an ML-based model is proposed to categorize valid versus illegitimate privacy policies. The related research is discussed in Section 5.6. Finally, Section 5.7 concludes the study and identifies future research issues.

5.2 Related Work

This section discusses associated privacy-related research issues. The existing literature can be grouped into three areas: personally controlled EHR systems, blockchain-based EHR systems, and context-sensitive privacy policies.

5.2.1 Personally-Controlled EHR Systems

Personal Electronic Health Record (PCEHR) systems enable individuals to manage their health information and control access. However, this also requires individuals to safeguard their data. Privacy is a crucial factor in sensitive sectors such as healthcare, and non-compliance can lead to substantial penalties. Unfortunately, many large health information systems still display privacy problems and identification risks for users due to inadequate implementation of legal requirements [398, 399]. Various proposals (e.g., [392, 397]) have been presented to address privacy concerns in personal health records, but they frequently lack empirical evidence and real-world tests and did not address potential ethical and legal concerns of implementing such systems [316]. Similarly, a proposed privacy-preserving personal health record (P3HR) system lacked a comprehensive security and performance evaluation [390], while a proposed Hippocratic database approach did not provide empirical evidence or case studies to support its efficacy [401]. An essential aspect to consider when developing personally controlled EHR systems is striking a balance between privacy and accessibility [6]. It is critical to protect the privacy of patient health information while ensuring that authorized healthcare providers can access the information they need to provide effective care. Another important factor when developing personally controlled EHR systems is to ensure that they are user-friendly and accessible to all patients [388, 389], regardless of age, education, or technological knowledge. This is challenging due to the complex nature of health information and the variety of devices and platforms used to access EHR systems. Mamum et al. [395] proposed a homomorphic encryption approach to encrypt patient information. The decryption key will be used by the patient, ensuring that no other person can access your information without prior authorization. To improve reliability and privacy, a cryptographic verification technique is introduced to ensure that only the authorized person has access to the corresponding records [396].

Privacy is a vital factor in sectors such as healthcare, banking, and defense, where sensitive and confidential data must be protected from unauthorized parties [72]. Numerous legislative rules and regulations have been introduced in European countries to ensure citizens' privacy [72, 387]. Global data protection standards have been established, which

outline specific data protection requirements and noncompliance penalties. According to Baker [398], patient care involves providing relevant care to individual patients according to their preferences, needs, and values, and ensuring that good clinical decisions are made. This patient care includes involving, informing, and listening to patients. Due to recent digital transformations in the healthcare sectors and associated data and privacy breaches, rebuilding trust in health-related IT systems has become an urgent challenge.

While personally controlled EHR systems have the potential to enhance privacy and patient empowerment in healthcare, several challenges must be addressed to ensure their effectiveness and acceptability. These challenges include balancing privacy and accessibility, making EHR systems more user-friendly and accessible, and acknowledging the cultural and social context of EHR system development and implementation. Overcoming these challenges will require further research, collaboration and innovation between healthcare providers, researchers, and technology developers.

5.2.2 Ensuring Privacy through Smart Contract—Healthcare Blockchain Systems

Blockchain-based EHR systems are increasingly gaining recognition for their potential to enhance security and privacy in the management of health data. By leveraging distributed ledger technology, these systems can effectively prevent unauthorized access and data breaches. However, challenges still need to be addressed when implementing blockchain systems in healthcare, particularly when sharing patient information with multiple stakeholders.

Recent studies have explored the use of blockchain technology to improve security and privacy in healthcare IT systems. In [402], the authors proposed a consortium blockchain for secure and privacy-preserving data sharing in e-health systems. Although their study provided an in-depth description of the proposed architecture and its benefits, it lacked empirical evidence and real-world evaluations and did not discuss potential limitations or challenges associated with implementing such a system. In [403], the study examined the applications of distributed ledger technologies of blockchain in biomedical and healthcare settings [403]. Although the authors thoroughly reviewed the existing literature and proposed various use cases, the study was published in 2017, and blockchain technology has evolved significantly since then. Furthermore, the authors did not address the potential drawbacks or limitations of using blockchain in healthcare settings. In [404], the authors focused on the potential of blockchain technology to improve the security and privacy of healthcare data stored in the cloud. The authors provided a comprehensive overview of the challenges and explained how the blockchain could address them. However, the

article did not critically evaluate the technology's limitations and challenges, such as scalability and interoperability issues. In [487], the authors proposed a blockchain-based incentive mechanism for crowd-sensing applications that preserve privacy. Although presenting an interesting idea, the study lacks sufficient detail on technical implementation and evaluation and does not compare the proposed mechanism to existing solutions or discuss limitations or future work. In [488], the authors introduced a blockchain-based solution called Medblock for the efficient and secure sharing of medical data. The authors claimed that their system could overcome traditional centralized data storage limitations, but they did not provide a comprehensive evaluation of the proposed system's scalability and efficiency or detailed information about its implementation. Finally, in [489], the authors proposed a blockchain healthcare system using smart contracts to secure automated remote patient monitoring. While the authors presented a detailed description of the proposed system and a theoretical analysis of its security and privacy features, they lacked empirical evidence to support the feasibility and effectiveness of the system and did not address potential challenges in implementing the system in a real-world healthcare setting.

Although blockchain-based EHR systems can offer significant benefits in terms of security and privacy, there are still challenges and limitations to be addressed, especially when sharing patient data with multiple stakeholders. More research, validation, and critical analysis is needed to ensure the practicality, scalability, and effectiveness of these systems in real-world healthcare scenarios.

5.2.3 Context-Sensitive Privacy Policies

In recent years, there has been a growing interest in context-sensitive approaches within the EHR domain. In [372], the study presented a context-aware access control model for cloud-based data resources, incorporating imprecise context information. The authors used fuzzy logic to model uncertainty in context information and developed a context-aware access control framework. However, the study did not comprehensively evaluate the proposed model, including a comparative analysis with other state-of-the-art approaches, scalability, and performance testing. Additionally, no practical implementation of the proposed framework in real-world settings was mentioned. Although the proposed approach seemed promising, the lack of evaluation and practical implementation made it difficult to assess its effectiveness and feasibility. In [386], the article introduced a policy model and framework for context-aware access control of information resources. Their model integrated contextual factors such as user identity, location, and time to determine access privileges. However, it lacked empirical validation of the proposed framework, leaving its effectiveness uncertain in real-world scenarios. In addition, the article did not

address the potential ethical implications of context-aware access control, such as privacy and discrimination concerns. Further research and analysis are required to address these issues. In [393], the article proposed a fog-based context-aware access control (CAC) system to achieve security scalability and flexibility. The authors argued that their system could improve security in fog computing environments by providing dynamic and context-aware access control. The article provided a comprehensive overview of the proposed CAC system and discussed its details of implementation. However, the article lacked empirical evaluation of the performance and scalability of the proposed system. Additionally, it did not address the potential challenges and limitations of implementing such a system in real-world scenarios. Overall, the proposed system appeared promising, but further research is needed to validate its effectiveness and practicality. In [490], the study suggested an ontology-based approach to dynamic contextual role-based access control in pervasive computing environments. The authors described the architecture of the proposed system and evaluated its effectiveness through simulations. However, the evaluation of the system was limited to simulations and a real-world implementation and evaluation of the approach would be advantageous. Additionally, the study could benefit from a more in-depth discussion of related work in the field of contextual role-based access control.

In summary, while these context-sensitive approaches have made strides in proposing enhanced protection for EHRs, they have proven insufficient to accurately model relevant stakeholders and health information.

5.2.4 Homomorphic Encryption in EHR Systems

The role of homomorphic encryption in preserving the privacy of EHRs has been explored in various studies, which have claimed that the approach offers computation on encrypted data without requiring decryption, effectively facilitating secure data sharing and collaboration. Paul et al. [491] constructed a privacy-preserving framework, using homomorphic encryption to protect EHRs during collaborative machine learning processes. Although the proposed framework had potential, the study did not adequately address the limitations of the framework, including potential vulnerabilities of the encryption scheme, scalability, and maintaining confidentiality during collective learning. Ikuomola et al. [492] addressed privacy concerns in e-health clouds using homomorphic encryption and access control. However, the research was marked by the absence of a detailed analysis of the effectiveness of the solution. Furthermore, potential vulnerabilities or attacks that could undermine the security of the proposed system and scalability issues related to large-scale e-health cloud environments were not adequately addressed.

Vengadapurvaja et al. [493] developed an efficient homomorphic medical image encryption algorithm for secure medical image storage in the cloud. Despite its focus on medical images, the approach did not extend to the encryption of other types of EHR data. This narrow scope limited its comprehensive application to broader privacy concerns about EHRs. Alzubi et al. [494] integrated homomorphic encryption with deep neural networks to secure the transmission and diagnosis of medical data. However, unspecified inadequacies were identified in preserving the privacy of the EHR. A thorough examination of the study would provide a better understanding of these limitations. Subramaniaswamy et al. [495] implemented a somewhat homomorphic encryption scheme for edge devices based on IoT sensor signals. However, without detailed information from the study, it is difficult to identify specific inadequacies in the preservation of the privacy of the EHR. Potential challenges could include the scalability, performance or vulnerability of the implemented scheme when applied to real-world EHR systems. Finally, Vamsi et al. [496] investigated various homomorphic encryption schemes to protect EHR in the cloud environment. Despite potential benefits, several inadequacies in the application of homomorphic encryption to preserve EHR privacy were noted. Challenges, such as the overhead performance of homomorphic encryption, integration difficulties with existing healthcare systems, and the need for efficient key management strategies, were some of the identified concerns.

Although various studies have explored the role of homomorphic encryption in preserving the privacy of EHR, each presents certain inadequacies. The key among these is the vulnerability of the encryption schemes employed, limitations in scalability, difficulties in maintaining the confidentiality of sensitive data, and the substantial computational overhead that their encryption techniques have introduced. Additionally, a narrow focus on specific data types, such as medical images, excludes comprehensive coverage of EHR privacy concerns. The challenges in integrating homomorphic encryption schemes into existing healthcare systems, including the issues of interoperability, data access control, and key management strategies, further compound the problem. This study aimed to address these shortcomings by proposing a novel privacy-preserving approach for EHRs, taking advantage of the benefits of homomorphic encryption while addressing its limitations. We sought to develop a robust, scalable and versatile homomorphic encryption scheme that can protect various types of EHR data. Our methodology focused on ensuring efficient performance, facilitating secure data sharing, and improving integration with existing healthcare systems. Furthermore, we will offer solutions for effective key management, ensuring a holistic and comprehensive approach to preserving EHR privacy.

5.2.5 Comparison with Our Study

Our study acknowledged that several attempts have been made to address privacy and security issues within the realm of EHRs. While crucial, these efforts often exhibit certain shortcomings. For example, they often do not consider concerns raised by previous research and lack comprehensive and robust evaluations of their effectiveness, scalability, and process efficiency. In addition, these studies sometimes focus too narrowly on specific solutions, such as homomorphic encryption or context-sensitive privacy policies, overlooking the need for more holistic and comprehensive strategies that can navigate the complexities of modern healthcare systems. Furthermore, the scalability of these solutions, especially when implemented in larger and more diverse healthcare systems, often remains inadequately explored. Another underaddressed concern pertains to the process efficiency of these proposed solutions. In the fast-paced, high-stakes environment of healthcare, solutions that are computationally demanding or overly complex may not be feasible, despite their theoretical advantages. Our added perspective does not devalue the existing body of work. Instead, it illuminates the multifaceted nature of privacy and security in healthcare data management. Our research aspired to address these challenges through an approach that balances privacy protection, process efficiency, scalability, and real-world applicability. We aimed to build on these prior efforts, incorporating their strengths and also striving to rectify their shortcomings.

With the aim of applying an ontology- and ML-based approach to protect health information, our study sought to explore new solutions to challenges in this domain. The distinguishing aspects of our research are as follows:

- ✓ **An Attempt at Comprehensive Privacy Protection:** Our approach endeavored to create a privacy protection solution that is more robust and specific to healthcare information systems by combining ontology and ML. Although we believe that it can offer improved protection, further studies and real-world testing are necessary to validate this claim.
- ✓ **Consideration of Legal Requirements:** We made an effort to incorporate the breadth of GDPR and other privacy regulations into our proposed solution, aiming to ensure compliance with the necessary legal requirements. However, adapting to evolving legal landscapes will require continuous updates and adjustments.
- ✓ **Exploration of Balancing Privacy and Accessibility:** Our proposed solution attempts to balance the preservation of patients' health information privacy with the necessity of access for authorized healthcare providers. Future studies should focus on how well we have achieved this balance in various real-world scenarios.

-
- ✓ **Aim for User-friendly and Accessible Systems:** We recognize the complex nature of health information and the diverse range of devices and platforms used to access EHR systems. Our study aimed for a more inclusive approach to healthcare information management, although the user-friendliness of our solution has not yet been evaluated in diverse groups of patients.
 - ✓ **Emphasis on Real-world Implementation and Evaluation:** We aimed to provide solutions that can be implemented and evaluated in real-world healthcare settings. Our approach takes a practical perspective, although extensive empirical evidence to support its effectiveness has yet to be collected.
 - ✓ **Acknowledgment of the Cultural and Social Context:** In our study, we considered the cultural and social context of the development and implementation of the EHR system, with the goal of finding a solution that can accommodate diverse needs. However, more research is required to confirm the adaptability of our solution to different cultural and social contexts.

Our study offers an exploration of the combination of ontology and ML to protect privacy in healthcare information systems. By acknowledging the importance of factors such as legal compliance, balancing privacy and accessibility, creating user-friendly systems, real-world implementation, and considering cultural and societal aspects, we strove to extend the knowledge in the field. However, it is important to note that our proposed solution is a preliminary attempt, and further validation through future research is needed. The potential impact of our study is in providing new perspectives and suggesting areas of focus for ongoing exploration in the field of healthcare information management.

5.3 Research Motivation

In this section, we dive into a variety of application scenarios, examine them, and highlight research challenges that remain unaddressed in the current EHR privacy literature. Today, privacy is a vital concern in cybersecurity, and protecting patient data is essential by implementing robust EHR privacy and security policies on both national and international levels [382].

To further strengthen our motivation for this research, we reflect on real-world examples that underline the privacy and security challenges currently plaguing the realm of health information sharing. For example, the considerable data breach at Anthem Inc. In 2015, where hackers gained unauthorized access to the personal information of nearly 78.8 million individuals [497], showcases the vulnerabilities of large-scale health data

systems are highlighted. Despite the robust security measures in place, the attackers were able to exploit weak points in Anthem's system, leading to a catastrophic loss of privacy for millions of people. Such breaches clarify the crucial need for improved privacy protection mechanisms, specifically those that are capable of safeguarding Personally Identifiable Information (PII) against increasingly sophisticated forms of cyberattacks.

Furthermore, real-life examples can also provide insight into the effectiveness of existing privacy regulations in the face of evolving technological landscapes. For example, consider the case of the UK's National Health Service (NHS) in 2018, when it was discovered that third party organizations were purchasing anonymized patient data for market research [498]. Despite adhering to existing privacy regulations, the anonymization techniques employed failed to prevent the reidentification of individual patients from the purchased data, leading to serious privacy concerns. These incidents not only demonstrate the importance of our research but also highlight the urgent need for an integrative approach that combines machine learning with ontologies to secure EHRs effectively. Our proposal aimed to identify and protect PII within privacy policies, adding an additional layer of security to existing frameworks. Such a solution can potentially prevent future privacy violations, particularly those related to the reidentification of anonymized data, thereby ensuring the integrity of patient information in digital health platforms.

An appropriate privacy model is needed to allow patients to have control over their own data, and it could also facilitate tracking of who has accessed their health information and the parties to whom it has been sent [499]. As mandated by the Australian Privacy Principles (APP) 1988 (<https://www.oaic.gov.au/privacy/australian-privacy-principles> (accessed on 6 July 2023)), patients should be informed about the data collected and the way in which their personal health information is used. During visits to hospitals or clinics, patients should also be notified of the reasons behind collecting and using their data, the duration of data retention, and the parties with whom they will be shared. According to NHS England, centralizing health information at a national level is crucial. When a General Practitioner updates a patient's registration information in their clinical system, the Primary Care Support England (PCSE) system leverages this information to update the National Health Application and Infrastructure Services (NHAIS), responsible for maintaining the National Patient Register. The Royal Australian College of General Practitioners (RACGP) has also introduced a sample registration form for new patients [383]. To comply with federal and state privacy laws, this form aligns with the RACGP standards for general practices. If patients have privacy concerns, they can discuss them with their GP and opt to leave the form blank. However, it is not considered best practice to let patients leave the form blank, as the information may be crucial at any stage of their treatment, and a lack of data could result in improper treatments.

Various individuals or groups, including health-related and non-health stakeholders, can engage with EHR systems. Stakeholders can be categorized and arranged differently based on their role in the management of EHR records (Table 5.1). In the following paragraphs, we will explore several such scenarios.

TABLE 5.1: List of role-based stakeholders and privacy rules.

| Stakeholders' Category | Stakeholders' Example | Role-Based (Senior/Junior) | Privacy Roles |
|-------------------------------|----------------------------------|-----------------------------------|---|
| Support Professionals | Receptionist, Pharmacist | Junior | Policy 1: Support professionals can only deal with personal health information |
| Nursing Professionals | Nurse Manager, Nurse | Junior | Policy 2: Nursing professionals can deal with personal and private health information |
| Medical Practitioners | General Practitioner, Specialist | Senior | Policy 3: Medical practitioners can deal with personal, private, sensitive, and historical health information |
| Diagnosis Professionals | Radiologist, Medical Technician | Senior | Policy 4: Diagnosis professionals can deal with personal and historical health information |
| Medical Scientists | Researcher, Junior Researcher | Junior | Policy 5: Medical scientists can deal with all types of health information with the approval of relevant stakeholders |

5.3.1 Use Case Scenarios

In order to establish a comprehensive privacy ontology, current scenarios must be improved and diversified, taking into account the multifaceted reality of healthcare organizations. The following use case scenarios span a range of situations, each with differing stakeholders, types of health information, and privacy concerns. These cases, although varied, represent a snapshot of the highly complex landscape of privacy preservation in the context of EHRs.

5.3.1.1 Scenario 1: Primary Care Physician

An elderly woman, living with her oldest son, is struggling with her mental health after witnessing the sudden death of her youngest grandson. As she shows signs of distress and paranoia, her son seeks the help of a primary care physician (PCP). After an examination, the PCP suggests consulting with a mental health professional. In this scenario, privacy concerns relate to the sensitive nature of the woman's mental health status and her medical history.

- Stakeholders: PCP, patient, patient's son, and mental health professional.
- Health Information: Private (address, location), Personal (demographic details), Sensitive (medical and psychological history), Historical (previous health evaluations).

5.3.1.2 Scenario 2: Emergency Care

After a serious car accident, an unconscious patient is rushed to the Emergency Department (ED), where a nurse evaluates him. The patient's Medicare details are used to gain access to his medical history to determine the best course of urgent care. The privacy concern here relates to the patient's inability to provide consent for access to his medical records due to his unconscious state.

- Stakeholders: Nurse, patient, medical team, and Medicare.
- Health Information: Private (address, location), Personal (demographic details, Medicare details), Sensitive (medical history, accident details).

5.3.1.3 Scenario 3: Clinical Research

A breast cancer patient undergoing radiotherapy expresses concerns to her GP about her family history, particularly since her mother died of brain hemorrhage. Her GP consults with a research team to access clinical trials data and explore the prevalence of similar cases. The patient's personal information, medical history and family history should be handled discreetly due to the sensitive nature of her condition and personal fears.

- Stakeholders: Researcher, GP, patient, clinical trials team.
- Health Information: Private (address, location), Personal (demographic details), Sensitive (medical history, family history), Historical (previous treatment records).

5.3.1.4 Scenario 4: Multidisciplinary Consultation

A patient with a rare genetic disorder requires consultation with a multidisciplinary team, involving primary care physicians, specialists, therapists and social workers. The complexity of the case necessitates sharing extensive patient data across the team while ensuring the patient's privacy.

- Stakeholders: Primary care physicians, specialists, therapists, social workers, patient.
- Health Information: Private (address, location), Personal (demographic details), Sensitive (medical and genetic history), Historical (treatment and therapy records).

5.3.1.5 Scenario 5: Telehealth

A remote patient receives care via a telehealth platform. Patients' electronic health records should be accessed and updated by healthcare providers during virtual consultations. Privacy concerns here arise due to the potential vulnerabilities associated with the transfer of sensitive health data over digital channels.

- Stakeholders: Patient, healthcare providers, provider of the telehealth platform.
- Health Information: Private (address, location), Personal (demographic details), Sensitive (medical history), Historical (previous consultation records).

5.3.1.6 Scenario 6: Data Breach

A healthcare organization experiences a data breach and the EHRs of multiple patients are potentially compromised. This scenario raises significant privacy concerns related to unauthorized access and potential misuse of health data.

- Stakeholders: Patients, healthcare organization, IT department, potentially unauthorized third parties.
- Health Information: Private (address, location), Personal (demographic details), Sensitive (medical history), Historical (previous treatment records).

These enhanced scenarios provide a broader understanding of the complex landscape of privacy preservation in EHRs and should provide a strong foundation for the development

of a comprehensive privacy ontology. The complexity and diversity of these scenarios reflect the dynamic nature of healthcare care delivery and the myriad of privacy concerns that arise in real-world healthcare settings.

5.3.2 Research Challenges

In the context of the GP and researcher scenarios previously discussed, we have recognized several challenges that must be tackled when developing a universal privacy ontology for various EHR platforms, such as MHR and NHS systems. Addressing these challenges is crucial to ensure the privacy and security of sensitive patient information while enabling seamless data exchange across different EHR platforms.

1. Sharing personal information with relevant stakeholders: According to APP 6 [482], personal information should only be shared with relevant stakeholders for a specified purpose, such as treatment or daily care with GPs or nurses. Personal health information may also be shared for secondary purposes under certain conditions. However, perceiving these secondary purposes within the existing EHR literature proves to be a substantial research challenge. This difficulty also extends to sensitive health information. Although the principles of APP are generic and can be applied to any domain, implementing these principles within the health information domain is particularly challenging. As a result, the development of a privacy ontology for EHR systems is necessary.
2. Identifying relevant stakeholders: One of the detailed challenges involved in building the privacy ontology includes identifying the relevant stakeholders associated with different EHR systems. These stakeholders can range from healthcare providers, insurance companies, and government agencies to patients themselves. A privacy ontology should be able to accommodate these various stakeholder groups and their respective access levels, ensuring that sensitive patient information is only accessible to those with appropriate authorization.
3. Categorizing different levels of health-related patient information: Another challenge is to categorize different levels of health-related patient information, which can range from general health indicators to highly sensitive data, such as genetic test results or mental health records. Creating a privacy ontology that can effectively classify this information is essential for implementing appropriate access controls and maintaining patient confidentiality.
4. Definition of privacy rules and policies: The privacy ontology should also define privacy rules and policies for relevant health-related stakeholders, allowing them

to share patient health records across different EHR platforms securely. These rules and policies should be robust, flexible, and adaptable to meet the diverse and evolving needs of different healthcare systems and their stakeholders.

By addressing these challenges, a universal privacy ontology can be developed for EHR platforms, providing a framework to ensure the privacy and security of patient information while facilitating interoperability and collaboration among healthcare stakeholders. This ontology will ultimately improve the efficiency and effectiveness of healthcare care delivery, which will benefit both patients and providers.

5.4 A Privacy Model for EHR Systems

This section introduces a privacy model for EHR systems to address the research challenges identified in Section 5.2.

5.4.1 Leveraging Ontology and ML for Enhanced e-Healthcare Privacy

Recent years have witnessed a growing emphasis on the connection between privacy ontology and ML in the context of e-Healthcare systems. These approaches have been used to improve various aspects of healthcare care systems, such as intrusion detection, confidentiality, and privacy of EHR. In this section, we will discuss the key themes surrounding ontology and ML in e-Healthcare systems.

5.4.1.1 Intrusion Detection and Prevention

Sreejith and Senthil's research [500] proposed a model to detect intrusion attacks based on a NoSQL database and semantic features. This model highlights the role of ML in detecting and preventing real-time intrusion attacks in healthcare systems.

5.4.1.2 Confidentiality and Privacy of EHR

In [501], their research focused on an ontological framework designed to improve the confidentiality and privacy of the EHR. Their framework aimed to detect anomalies in abnormal patterns of access to healthcare records, predict vulnerable healthcare records for prioritized security efforts, and analyze stakeholders' behavior to detect suspicious activity.

5.4.1.3 Improved Indexing and Retrieval Performance

ML and ontology-based techniques have been shown to enhance the effectiveness of indexing processes and retrieval performance in various studies [502, 503]. For example, a framework for smart e-healthcare systems employs IoT technology while maintaining privacy and authentication through a combination of encryption, secure authentication protocols, and Blockchain technology.

5.4.1.4 Secure Data Access and Privacy Preservation

Sun et al. [504] explored a bilateral fine-grained access control mechanism in cloud-enabled industrial IoT for healthcare care, using Blockchain-based frameworks for granular access control, secure data access, and privacy preservation.

5.4.1.5 Privacy Disclosure Measurement

Research on ontology-based approaches to protecting personal information in online privacy policies suggests that these models offer a standardized and objective way to measure privacy disclosure [505]. Privacy-preserving ontology is analyzed through various stages, including data collection, data publication, and output with respect to modeling and training.

5.4.1.6 Efficiency and Accuracy in Clinical Information Extraction

Studies such as Yehia et al.'s [506] have demonstrated increased efficiency and precision in extracting clinical information from free text notes written by physicians using ML-based approaches.

5.4.1.7 Structured Approach to organizing Clinical Data

Bosco et al. [507] illustrated that ontologies can provide a structured and standardized approach to the organization of clinical data, supporting the interoperability between EHR systems, and ultimately improving patient care and facilitating clinical research.

To summarize, incorporating ontology-driven strategies alongside machine learning in EHRs yields considerable benefits in terms of privacy, security, and efficacy. The use of

ontology-focused techniques allows the creation of a standardized and structured framework to organize and handle health data, leading to enhanced interoperability, information exchange, and decision-making. Our research distinguishes itself by tackling various obstacles such as intrusion detection, EHR confidentiality, optimized indexing and retrieval performance, secure data accessibility, privacy conservation, and clinical data extraction. This comprehensive approach guarantees the protection of sensitive health data and empowers healthcare professionals to provide superior patient care.

By persistently examining and refining these methodologies, we can further advance electronic healthcare systems and contribute to the creation of more secure, privacy-focused, and effective solutions to manage delicate health information. Consequently, our ontology-centered method, in conjunction with machine learning tactics, possesses immense potential to secure health data and improve overall patient care results.

5.4.2 Conceptual Privacy Models

We identified the concepts behind the privacy model for different EHR systems. A brief description of the concepts is presented in Figure 5.1, which can be titled the identification layer. It shows the different types of stakeholders (e.g., GP, nurse, policy maker), technologies (e.g., the digital platforms to interact with relevant stakeholders), and health information (e.g., personal, sensitive) involved in the EHR scenarios.

- Identify stakeholders: the relevant stakeholders need to be identified from the application scenarios.
- Identify technologies: the relevant health technologies and platforms need to be identified to gather health related information.
- Identify health information: Health information and records need to be identified and labeled in different categories, such as private record, sensitive record, etc.

5.4.3 Identifying Stakeholders

We Analyzed various EHR scenarios, including those for receptionists, nurses, policy makers, and media personnel, as well as those outlined in Section 5.4.2. Based on our findings, we have identified several stakeholders associated with EHR platforms, which we have listed in Table 5.1.

To simplify the privacy model and utilize the inheritance concept of context-aware role-based access control systems (as described in [385, 393]), we classified EHR stakeholders

into two categories: primary and secondary stakeholders. Primary stakeholders, such as GPs and nurses, are directly involved with EHR platforms and have access to patient health information. Secondary stakeholders, such as media personnel, use health information without being directly involved with EHR systems. We also defined two types of roles for different stakeholders to create a privacy model applicable across different stakeholder types. These roles include senior roles and junior roles. Table 5.1 shows the inheritance relationship of the role between stakeholders, where top-level stakeholders, such as medical practitioners and medical scientists, are classified as senior roles, while bottom-level stakeholders, such as GPs and researchers, are classified as junior roles. This approach helps ensure that access to sensitive patient information is appropriately controlled and that privacy is maintained for all stakeholders involved with EHR systems.

5.4.4 Redefining Health Information and Privacy Rules

In response to valuable feedback and mindful of evolving privacy legislation such as the GDPR and several national laws, we revisited our initial classification of health information and redefined them as follows:

- **Identifiable Health Information:** This includes any data that can be used to identify an individual, such as name, address, and location. This is similar to what was previously described as 'private health information' and 'personal information.' Under GDPR and similar laws, all health information is considered sensitive, hence our shift towards a unified category.
- **Health Record Information:** This encompasses the clinical details of a patient's health condition and medical history. It includes past diagnoses, treatment records, and other medical reports. This category is more compatible with current legislation and consolidates what we previously classified as 'sensitive information' and 'historical information'. The important distinction here is that this information is sensitive by its very nature and is treated as such under GDPR and similar laws.

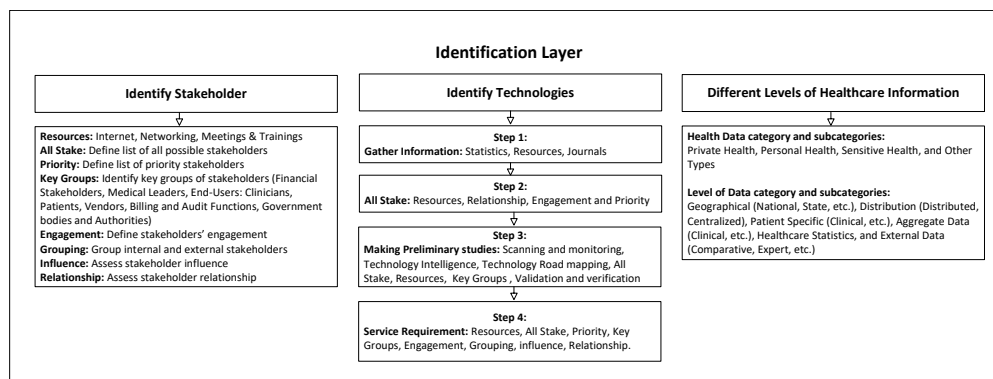


FIGURE 5.1: The relevant concepts to build the privacy model.

With this redefinition, we present an updated privacy ontology and its associated concepts, including a role ontology, health information ontology, and privacy policy ontology. These concepts, which are based on the revised privacy model introduced earlier, can be viewed in Figures 5.2–5.5. These modifications ensure that our ontology is in line with current legislation, making it more applicable to complex healthcare and health-related information landscapes.

Our ontology and knowledge bases were defined using the widely used Web Ontology Language (OWL) [508], specifically Prot'eg'e OWL 5.5 (<https://protege.stanford.edu> (accessed on 6 July 2023)). We employed an object-oriented approach to model various stakeholders, health information, and privacy rules. It includes classes, subclasses, datatypes, and object type properties. The health-related stakeholders were defined as classes and sub classes (i.e., primary and secondary stakeholders) and their relevant properties were defined using class-to-class object type and datatype properties.

Table 5.2 provides a technical description of our updated privacy ontology. It comprises three core concepts: role ontology, health information ontology, and privacy policy ontology. This revised framework reflects a more accurate portrayal of the current healthcare landscape and the information contained within the EHRs, making it more suitable for privacy preservation.

TABLE 5.2: The modeling criteria of the ontology.

| Basic Modeling Criteria | Privacy Ontology Elements |
|-------------------------|---|
| Classes | All primary stakeholders are represented as senior roles. |
| Sub classes | All secondary stakeholders are represented as junior roles. |
| Object-type Properties | The relationships between classes to classes are represented as object-type properties. |
| Datatype Properties | The relationships between classes to their features are represented as datatype properties. |

Our updated ontology provides a better foundation to deal with complex healthcare and health-related information environments. Identifiable health information and health record information better adhere to GDPR and other national laws, providing a robust framework for data protection. With this new approach, the ontology can better reflect the complex and diverse processes and results found in healthcare systems, ensuring that the EHRs remain secure and the privacy of patients is respected.

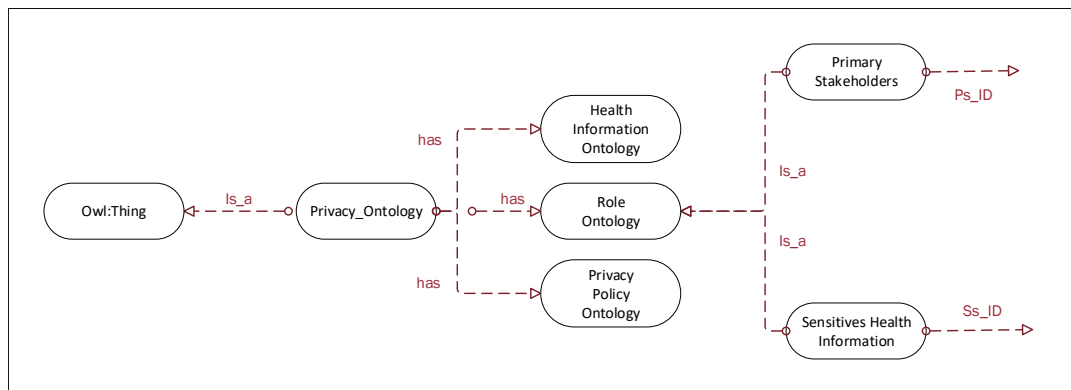


FIGURE 5.2: The core concepts of privacy ontology.

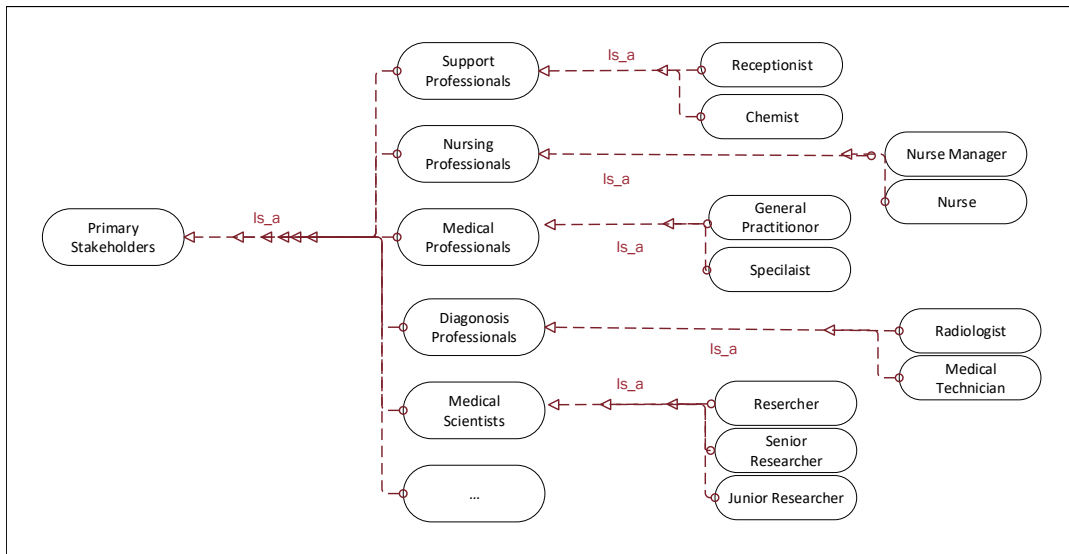


FIGURE 5.3: Role ontology.

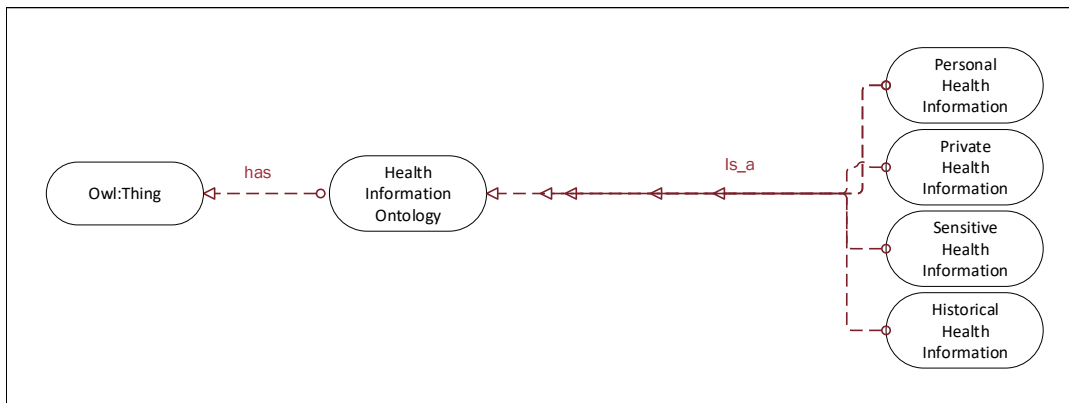


FIGURE 5.4: Health information ontology.

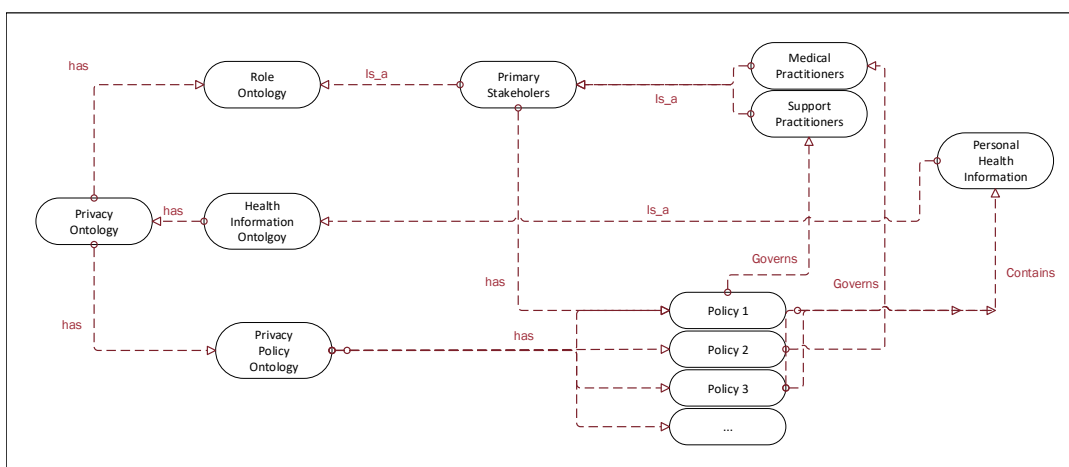


FIGURE 5.5: Privacy policy ontology.

5.4.5 Role Ontology

In this section, we will introduce a snapshot of the role ontology, which is based on the different stakeholders associated with EHR environments. Figure 5.2 shows the core concepts of the privacy ontology, which consists of three parts: Role Ontology, Health Information Ontology, and Privacy Policy Ontology. Primary Stakeholder and Secondary Stakeholder classes are both subclasses of the Role Ontology class. The “Is $_a$ ” property indicates the relationship between the Role Ontology class and its sub classes. Each class and subclass are defined by its datatype properties, such as the Primary Stakeholder class having a datatype property “Ps_ID”.

Primary stakeholders, such as healthcare professionals, can directly access patients’ health information through EHR systems, whereas secondary stakeholders, such as media professionals, can use some health information without being directly involved with EHR systems. Figure 5.3 provides a snapshot of the primary stakeholders. For the purpose of this chapter, the Role Ontology (see Figure 5.3) includes five domain-specific key classes: Support professionals, nurses, medical professionals, diagnostic professionals and medical scientists. Each class has its own set of sub classes; for example, the Nursing Professionals class has Nurse Manager and Nurse sub-classes. The proposed privacy ontology can be expanded to include new classes and sub-classes.

Understanding the integral role of ontology in improving data privacy in EHRs is paramount. In our proposed Role Ontology, classes such as ‘Primary Stakeholders and ‘Secondary Stakeholders’ explains who has authorized access to the EHRs and the extent of that access. In scenarios where a healthcare professional (primary stakeholder) accesses patient data directly from EHRs, or a media professional (secondary stakeholder) uses some health information without direct EHR involvement, it is the defined datatype properties such as ‘Ps_ID that regulate this access. Such a structured hierarchy of data access, based on well-defined classes and subclasses, ensures only authorized access to sensitive patient data, thereby enhancing data privacy. Additionally, our ontology model provides flexibility in expanding to include new classes and subclasses as required, ensuring the model’s scalability and adaptability to growing and diversifying healthcare data needs. Through our role ontology, we aimed to build a robust privacy-preserving framework where the right to access and the extent of that access is predefined based on the role of the stakeholder. This approach significantly mitigates unauthorized access, reduces privacy breaches, and promotes data confidentiality in EHR environments.

5.4.6 Health Information Ontology

In this section, we present an overview of the ontology of health information, which is based on various types of medical information found in EHR systems. Our privacy ontology proposal includes classes and subclasses of relevant health information. The core concept of the Health Information Ontology consists of domain-specific concepts such as Personal Health Information, Private Health Information, Sensitive Health Information, and Historical Health Information. The different types of health information are represented as sub-classes, and are linked to the core concept of the Health Information Ontology using an “Is_a” relationship, which is an object type property. You can see a snapshot of the health information ontology in Figure 5.4.

5.4.7 Privacy Policy Ontology

In this section, we look at the privacy policy ontology, which is based on the privacy rules we identified in the previous section. Figure 5.5 presents a snapshot of the privacy policy ontology, which also includes domain-specific concepts from EHR environments. To illustrate, according to Policy 1, Policy 2, and Policy 3 (as shown in Table 5.1), all physicians are authorized to access patients’ health information. Another example is that Policy 1 allows support professionals to access patients’ personal health information but not their entire medical records.

5.4.8 Disclosing Emergency Health Information for Patients in Car Accident Case Study

Case Study Overview: In the healthcare scenario we previously examined, nurses have the authority to access a patient’s personal health data, which they share with the main parties involved in the patient’s care. However, private information, such as the precise address of the patient’s home, remains confidential unless an urgent situation necessitates disclosure.

- Patient location: Somewhere, in a suburb of Melbourne.
- Primary stakeholders: Patient, GP, Paramedics, Emergency Room Nurse.
- Secondary stakeholders: Family members, Insurance provider.
- Patient Health Situation: A man with Type 2 diabetes living at 5 Somewhere Street in a Melbourne suburb regularly sees his GP for health monitoring and treatment.

His insurance provider and primary stakeholders have access to his health records, but not to his specific home address. One night, he is in a car accident and suffers from severe chest pain, difficulty breathing, and dizziness. A witness calls 000 for emergency help, and the operator dispatches paramedics after learning that the accident occurred in a Melbourne suburb.

Due to urgency, the operator shares the patient's exact address with paramedics, allowing them to reach him promptly and provide life-saving treatment. In emergencies, every second counts, and sharing the information of a patient can be vital for paramedics to act quickly, administer treatment, and transport the patient to a hospital if needed. Knowledge of the patient's home address helps plan the most efficient route to the hospital, which is crucial in some cases. Therefore, the emergency operator shares the address with the responding paramedics. On arrival, paramedics evaluate the patient's condition and suspect a heart attack. They provide oxygen and aspirin to stabilize him before transporting him to the closest hospital for further care. In the hospital, the emergency room nurse is informed about the patient's condition and gains access to his personal health data to aid in his treatment. In light of the emergency, the nurse also obtains the patient's private information, including his precise home address, to facilitate any required follow-up care or communication with family members.

Throughout the entire process, primary stakeholders, including the patient's GP, paramedics, and nurse, stay informed about the patient's condition and treatment. They work together to ensure that the patient receives the best possible care and support during this emergency. In this case, the patient's private health information is disclosed only to the primary stakeholders necessary in response to a critical emergency. This disclosure enables a fast and effective response that ultimately saves the patient's life.

A detailed health scenario (an emergency case is given in Table 5.3).

TABLE 5.3: A case study of a health emergency.

| Personal Health Information | Privacy Policy | Emergency Situation |
|--|--|---|
| A primary address, such as “suburb is a suburban area in Melbourne”, will only be released to all stakeholders. However, the patient’s actual address, which is “5 Somewhere Street, a suburban area in Melbourne”, is only released in the event of an emergency. | All primary stakeholders have access to patients’ personal health information. However, they do not have access to private health information. | In an emergency, some primary stakeholders can have access to patients’ private health information. For example, in cases where the patient is considered in a very critical condition and needs to be admitted to the hospital, physicians have access to patients’ exact home address for an emergency treatment/situation. |

5.5 Evaluation of Privacy Ontology and Experiments

We evaluated the privacy ontology by categorizing health-related privacy policies into two types: valid and illegitimate privacy policies. To differentiate between valid and illegitimate policies, we employed NLP-based ML models, specifically Bidirectional Encoder Representations from Transformers (BERT) [509]. Based on our proposed privacy ontology, we identified the following five steps to conduct the experiments.

- ✓ **Data Collection and Preprocessing:** We collected a large data set of health-related privacy policies from various online sources, such as hospitals, clinics, health insurance providers, and health-related mobile applications. Then, these policies were preprocessed to remove irrelevant information, formatting inconsistencies, and convert them into machine-readable format.
- ✓ **Annotation and labeling:** After preprocessing the data, we manually annotated and labeled the privacy policies based on their adherence to our privacy ontology. The annotation process involved experts in privacy and health domains, who categorized the policies into two classes: valid (policies that comply with the privacy ontology) and illegitimate (policies that do not comply with the privacy ontology).
- ✓ **Feature extraction:** We extracted relevant features from preprocessed privacy policies, such as the presence of specific keywords or phrases, using NLP techniques. These characteristics were critical for training the BERT model, which was then used to classify the policies into valid and illegitimate categories.

- ✓ **Training and Evaluation of the BERT Model:** We used the annotated and labeled data set to train the BERT model. During the training process, the model learns to identify patterns and relationships between the extracted features and the corresponding labels (valid or illegitimate). After training, we evaluated the model performance using standard evaluation metrics such as precision, recall, and F1-score [510].
- ✓ **Analyzing and Interpreting Results:** We analyze the results obtained from the BERT model to identify common patterns and trends in the classification of health-related privacy policies. This analysis provides valuable information on the effectiveness of the privacy ontology and helps identify potential areas for improvement.

Through these five steps, our aim was to demonstrate the effectiveness of our privacy ontology in distinguishing between valid and illegitimate health-related privacy policies. By leveraging state-of-the-art NLP techniques and ML models, such as BERT, we can automate the evaluation process, making it easier to ensure that privacy policies adhere to established privacy principles and guidelines.

5.5.1 Experiment Setup and Dataset Preparation

This section defines the processes involved in the preparation of the data set and the experimental setup to construct and evaluate our proposed privacy mechanism. To overcome previous limitations, our enhanced experiment design incorporated a broader range of healthcare privacy policies and adopts cross-validation techniques, ensuring a robust and comprehensive evaluation of machine learning models.

Our data set is an amalgamation of numerous health-related privacy policies, collected from various healthcare organizations worldwide. It includes 100 manually labeled policies categorized as valid privacy policies (labeled ‘1’) or illegitimate privacy policies (labelled ‘0’). Additionally, we incorporated 69 pre-labeled illegitimate privacy policies (labelled ‘0’) from the “SpywareGuide” online archive. The total dataset comprises 169 labelled privacy policies, providing a diverse and substantial foundation for training our privacy model. Labeling was performed according to specific terms of health privacy and security, in accordance with the APP. A custom Python script implemented on Google Colab facilitated the process, scanning for 13 specific phrases within each policy before labeling them as ‘1’ or ‘0’. This comprehensive 13-phase word analysis provided a rigorous method for categorization, ensuring accurate labeling.

Our research method approach involved using this data set in combination with four BERT-NLP techniques: BERT, Distil BERT, Albert Tokenizer, and Roberta Tokenizer.

The objective was to develop a machine learning-based privacy model specific to EHR environments, capable of accurately distinguishing between valid and illegitimate privacy policies. The model was built upon the principles derived from our updated privacy ontology, identifying a valid privacy policy as one that is either identical to or similar to a known policy previously labeled as ‘1’. In contrast, new policies that do not conform to these principles were considered illegitimate and labeled as ‘0’.

5.5.2 Clarifying Privacy Policy Classification

In our study, we used machine learning (ML) models to classify privacy policies as ‘valid’ or ‘illegitimate’. This raises the question: Why is there a need to classify a policy as legitimate or illegitimate, especially in a context where, as an example, healthcare providers in the U.S. must abide by the regulations established by the HIPAA [511], regardless of what the ML model dictates?

1. First, it is essential to note that, while healthcare institutions in the U.S. must follow HIPAA guidelines, not all institutions worldwide must adhere to the same guidelines. Different countries have different privacy regulations, and the level of privacy protection can vary significantly between countries and even across different institutions within the same country. This variability increases the importance of having a tool that can automatically assess the validity of privacy policies across a broad range of contexts.
2. Second, even within the U.S., not all healthcare entities are required to follow HIPAA regulations. HIPAA applies primarily to healthcare providers, health plans, and healthcare clearinghouses, but does not extend to entities such as life insurers, employers, or schools. Some of these entities might have access to sensitive health-related information and might formulate their privacy policies, necessitating a mechanism to evaluate the validity of their policies.
3. Third, even when healthcare entities are required to follow HIPAA guidelines, there may still be differences in how these guidelines are interpreted and implemented. Privacy policies can be complex and subtle, and different entities might have different interpretations of what constitutes a ‘valid’ policy under HIPAA guidelines. An ML model that classifies privacy policies can serve as an additional check on these entities’ interpretations, highlighting potential areas of concern.
4. Finally, while we used the example of HIPAA in our study, the principles of our privacy ontology and our ML-based privacy model are not limited to HIPAA. Our approach was designed to be adaptable to various privacy regulations, not just

HIPAA. This adaptability makes our approach potentially useful in a wide range of contexts, even in situations where the applicable privacy regulations differ from HIPAA.

Given these considerations, it becomes clear that there is a need to classify privacy policies as ‘valid’ or ‘illegitimate’ and that our ML-based privacy model can provide a valuable tool in this regard. It enables a more subtle understanding of privacy policies in different contexts, helping in the ongoing effort to ensure that sensitive health-related information is adequately protected.

5.5.3 Development of ML-Based Privacy Model

In this section, we present an improved ML-based privacy model to classify valid and illegitimate privacy policies in EHR environments. We have opted for four transformer-based text classification models, specifically BERT, DistilBERT, ALBERT, and RoBERTa, to categorize privacy policies. The rationale for choosing these models lies in their state-of-the-art performance in various natural language processing tasks, including text classification, sentiment analysis, and named entity recognition. Moreover, these models have demonstrated strong generalizability in different domains and languages.

The primary goal of our privacy model was to determine whether a privacy policy is ‘valid’ or ‘illegitimate’ based on the type and extent of personal, private, sensitive and historical information that the relevant organization usually collects from its clients (e.g., healthcare stakeholders requesting patient information). Our proposed privacy model employs an automated text classification mechanism to effectively distinguish between valid and illegitimate privacy policies.

To further enhance the model, we provide a detailed explanation of the four transformer-based text classification models.

- BERT: BERT is a powerful pretrained language model that has achieved state-of-the-art results in various NLP tasks. Its bidirectional nature enables it to understand the context from both the left and right sides of a word, resulting in improved understanding and classification accuracy.
- DistilBERT: DistilBERT is a distilled version of BERT that offers a smaller and faster alternative while maintaining most of the original model’s performance. This model is particularly useful in cases where computational resources or model size is a concern.

- ALBERT (A Lite BERT): ALBERT is another variant of BERT that reduces the size of the model and improves training efficiency by sharing parameters between layers. Despite its reduced size, ALBERT maintains competitive performance, making it a suitable choice for our privacy model.
- RoBERTa (Robustly Optimized BERT Pretraining Approach): RoBERTa is a modified version of BERT that has been optimized for increased training efficiency and performance. It uses dynamic masking, larger batch sizes, and other training optimizations to achieve superior results in text classification tasks.

Using these advanced transformer-based models, we have developed a machine learning-based privacy model derived from the concepts of our proposed privacy ontology. The objective of the model is to accurately and efficiently classify privacy policies in EHR environments as valid or illegitimate, ensuring the protection of sensitive patient information. We first manually annotated a training dataset of 169 policies, labeling each policy as valid (1) or illegitimate (0). A privacy policy was considered valid if it is identical or closely comparable to an already labeled “1” policy within the training data set. Policies that did not meet these criteria were labeled “0” and deemed illegitimate. An example of an illegitimate privacy policy sourced from the online data set “SpywareGuide.com” online dataset is shown in Table 5.4.

To enhance the performance of the model, we explored various BERT-based techniques to process the training data. These methods included transforming the data into tensors, creating data batches, fine-tuning the model, and finally validating its accuracy. We trained multiple variants of BERT and systematically compared their results to determine the most accurate model for our privacy policy classification task.

TABLE 5.4: An example privacy policy.

| Policy | Policy Statement | Type |
|--------|---|------|
| 0] | <p>... Alexa Internet Privacy Policy, based on the last updated version on 7 April 2011. What Personal Information About Users Does Alexa Gather? Information You Give Us—We receive and store any information you enter on our Web site or give us in any other way. Automatic Information We Collect from the Toolbar Service—When you use the Toolbar Service, we collect information about the websites you visit and the advertisements that you see on those websites, the searches you perform using search engines...</p> | |

5.5.4 Application of Privacy Ontology in the Machine Learning Experiment

In this section, we elaborate on how the privacy ontology model developed is integrated into the machine learning experiment and how it enhances the privacy of patient data. Our privacy ontology model is crucial in setting the ground rules for classifying privacy policies. It defines the valid and illegitimate privacy policy categories, shaping the direction of our machine learning model training. The privacy ontology model, constructed based on an in-depth analysis and understanding of privacy laws and principles, ensures a high standard in distinguishing valid and illegitimate privacy policies. The application of privacy ontology in our machine learning experiment can be elaborated in three distinct steps:

1. **Annotation and labelling:** The privacy ontology serves as a guide to annotating and labeling the collected privacy policies. The experts involved in the annotation process utilize the ontology's principles to identify the class of each policy. As such, the ontology enables a more reliable and consistent labeling process.
2. **Feature Extraction:** Privacy ontology plays a crucial role in determining the relevant characteristics of the privacy policies. It aids in identifying the specific keywords or phrases that signify valid or illegitimate policies, thereby helping in effective feature extraction.
3. **Model Training:** During the model training phase, the privacy ontology plays a critical role in guiding the learning process of the BERT model. The model learns to identify patterns that align with the principles of our privacy ontology, which helps to classify new privacy policies with precision.

Integrating our privacy ontology model into the machine learning experiment significantly enhances the privacy of patient data. It does so by ensuring that any privacy policy, whether from healthcare providers, health insurance providers, or health-related mobile applications, complies with established privacy principles and guidelines before being classified as valid. This procedure guarantees that only policies that uphold high privacy standards will be deemed valid, providing a robust safeguard for patient data privacy. It is worth noting that our ML model's potential to accurately classify privacy policies does not negate the human role in this process. Our model acts as an assisting tool that automates and speeds up the classification process, but the initial rules and principles (as defined by the privacy ontology) are still set by human experts. Thus, while our model provides an additional layer of protection for patient data privacy, it does not eliminate

the need for human oversight, particularly in complex cases where a human might be more adept at identifying the appropriate policy.

5.5.5 Evaluation Results

The main objective of this section is to identify legitimate privacy policies associated with health and personal information in different EHR environments. The primary objective was to ensure that the EHR remains secure from any unauthorized attempts to access it. Towards this goal, we conducted a set of experiments to evaluate the efficiency of the proposed ML-based privacy model. The technical details of the experiments and the relevant results have been presented below.

5.5.5.1 Technical Details

A valid privacy policy is represented as a 2-tuple relation, including privacy policies from our proposed privacy ontology and legitimate actions (e.g., collecting, storing and/or disseminating health information where personal, private, or sensitive information is involved with authorized parties). On the contrary, an illegitimate privacy policy is one in which illegitimate actions (for example, using health information for marketing purposes) are involved with unauthorized parties (e.g., who are not primary or secondary health related stakeholders).

We calculated accuracy, precision, recall, and the f1 score using different BERT techniques to classify privacy policies (the data set of 169 policies). These BERT models were used to model our ML-based privacy approach. We use an automated text categorization mechanism to classify a privacy policy as valid or illegitimate.

5.5.5.2 Dataset and Results

The proposed model was evaluated on a data set that contains valid and illegitimate privacy policies generated by health-related organizations. We collect those privacy policies from multiple health-related organizations. We also used some illegitimate privacy policies from the online SpywareGuide archive.

Table 5.5 shows the experiment's results using the different NLP-based BERT techniques, such as BERT, Distil BERT, Albert Tokenizer, and Roberta Tokenizer. In these experiments, we used the concept of automated text categorization mechanism in our ML-based privacy approach through our introduced privacy classification technique: illegitimate versus valid. We achieved 94% accuracy using the Distil BERT technique, which

is better than the other BERT techniques that achieved 76%, 90%, and 92% accuracy using Albert Tokenizer, Roberta Tokenizer, and BERT techniques, respectively.

In the aforementioned experiments conducted on Google Colab, a cloud-based machine learning platform, we used a custom dataset comprising privacy policies obtained from a variety of health-related organizations, both legitimate and illegitimate. Illegitimate policies were sourced primarily from the SpywareGuide archive, an online resource that provides information about privacy risks. Each privacy policy in our dataset was transformed into a vector representation using the BERT tokenizer before being processed by our machine learning model. For the experimental setup, we used the BERT, DistilBERT, Albert, and RoBERTa models as implemented in the Google Colab Transformers library. The models were fine-tuned on our data set using a learning rate of 2×10^{-5} , batch size of 16, and for a total of four epochs, leveraging the high-performance computing power of Google Colab. The choice of these parameter settings was informed by preliminary experiments and the recommended settings from the original studies that introduced these models. To validate the practicality of the proposed model, we performed an additional set of experiments using privacy policies from a diverse range of healthcare sectors, including hospitals, insurance providers, and digital health applications. We found that our model, trained and fine-tuned in Google Colab, consistently achieved high accuracy in identifying legitimate and illegitimate privacy policies across these sectors, reinforcing its practical applicability in a real-world context.

TABLE 5.5: Experiment results using NLP-based BERT techniques.

| BERT Techniques | Accuracy | Precision | F1 Score |
|----------------------------|-----------------|------------------|-----------------|
| BERT | 0.92 | 0.86 | 0.90 |
| Distil BERT | 0.94 | 0.94 | 0.94 |
| Albert Tokenizer | 0.76 | 0.87 | 0.76 |
| Roberta Tokenizer | 0.90 | 0.82 | 0.92 |

5.5.6 Summary of the Findings

In our experimental setup, we used various BERT techniques to distinguish between legitimate and illegitimate privacy policies, with the primary aim of establishing privacy measures that effectively protect the EHR from unauthorized access. To conduct this experiment, we used a data set consisting of 169 valid and illegitimate privacy policies collected from the online SpywareGuide archive and relevant health-related organizations. Our experiments demonstrated that the proposed ML-based privacy approach

can reliably recognize illegitimate policies, achieving an F1 score exceeding 0.94 when using a data test set comprising 20% of the data and a training dataset of 80%, indicating that our approach has been effective in identifying privacy policies that may not provide adequate protection for sensitive health information. The results of our analysis showed that Distil BERT outperformed the other techniques in terms of precision, precision, and F1 score, achieving balanced performance in identifying legitimate and illegitimate privacy policies. This finding suggested that Distil BERT may be an effective tool for analyzing privacy policies in eHealthcare systems, providing valuable insights for enhancing privacy protection measures.

We anticipate that the precision and precision of our approach could be further improved by using a larger dataset containing a more diverse range of health-related privacy policies. In future experiments, we plan to expand our data set and refine our proposed ML-based approach to better determine the validity of these privacy policies. By doing so, we hope to develop a more robust and accurate system to identify and protect against potential risks associated with inadequate privacy policies in electronic health systems.

5.5.7 Refined Technical Approach and Dataset Overview

This subsection delves into the sophisticated methodologies employed and offers a detailed exposition of the data set utilized in our experimentation, aiming to bolster the transparency and reproducibility of our findings.

5.5.7.1 In-depth Technical Methodology

The cornerstone of our approach is the use of advanced NLP models, with a particular emphasis on the variations of BERT including BERT [512–515], DistilBERT [60, 516–519], ALBERT [520, 521], and RoBERTa [522, 523]. These models were fine-tuned on a meticulously curated dataset of healthcare privacy policies. This process aimed at achieving a binary classification: categorizing policies into compliant (‘legal’) versus noncompliant (‘illegal’) with respect to prevailing privacy regulations such as HIPAA and GDPR.

The model fine-tuning adhered to a rigorously defined parameter set: a learning rate of 2×10^{-5} , a batch size of 32, over four training epochs. This parameterization emerged from an extensive series of preliminary trials, demonstrating an optimal compromise between computational efficiency and model precision [524–528].

5.5.7.2 Comprehensive Dataset Description

Our study employed a dataset composed of 169 distinct privacy policies, sourced from a broad spectrum of healthcare entities. Each policy underwent a meticulous labeling process, being categorized as ‘legal’ or ‘illegal’, grounded on its adherence to established privacy frameworks like HIPAA for U.S. entities and GDPR for European counterparts.

The dataset’s compilation was driven by the objective of encapsulating a broad array of healthcare sectors, thereby ensuring its representativeness. This breadth encompassed hospitals, insurance providers, and an array of digital health platforms. Subsequently, the dataset was partitioned into a training subset, constituting 80% of the total, and a testing subset, making up the remaining 20%, to facilitate a comprehensive evaluation of the model’s predictive prowess [524–528].

5.5.7.3 Ensuring Reproducibility

A commitment to enhancing the reproducibility of our results underpins this study. To this end, we pledge to provide access to both the utilized code and dataset upon formal request. This gesture aims to empower fellow researchers to either replicate our study or extend the dataset for further inquiry.

Complementarily, extensive documentation detailing the experimental setup—including software configurations, model parameters, and data preprocessing steps—accompanies our study. This documentation is crafted with the intention of enabling seamless replication of our experiments or the application of our methodology to novel datasets [520, 521, 529, 530].

To conclude, through elucidating our methodological rigor and providing a granular view of our dataset, we aspire to address the intricacies involved in our study. Our endeavor to make our resources accessible is driven by a commitment to foster reliability and encourage scholarly engagement within the domain of e-Healthcare privacy.

5.6 Discussion

This section will focus on discussing the key insights collected during the research.

5.6.1 The Relationship between the Ontology and the ML Model

Integrating an ontology-driven approach with ML could enhance the confidentiality of MHRs and refine data categorisation processes [531–536]. Primarily, an ontology-driven methodology offers a structured classification of medical terms and concepts. This structure facilitates the accurate identification and categorization of sensitive health data [535]. An organized and standardized approach to health information ensures interoperability and streamlines data management. Furthermore, the ontology can link various data components, delivering contextual information that could refine the performance of the ML algorithm when detecting sensitive information [533].

Training ML algorithms with a data set that contains pre-identified sensitive information can further enhance the accuracy of data categorization [532]. Here, the algorithm learns to recognize patterns and characteristics associated with sensitive data, ensuring the effective detection and protection of such data. In conclusion, combining an ontology-driven approach with ML offers a robust platform to protect medical records. This combination substantially improves the confidentiality and security of health data, contributing to better patient care and privacy.

5.6.2 The Complementarity of Ontology and ML Model

Our research illustrates that the interplay between ontology and ML can significantly enhance healthcare data management. The ontology's role is multifold: it serves as a semantic framework that provides context and meaning to raw data, it defines the scope of data to be processed, and it structures the data in a manner that can be effectively utilized by ML models [537, 538]. The ontology model plays an instrumental role in the preprocessing stage of ML by identifying relevant data sources and features [539]. This identification ensures that ML models are trained on relevant and meaningful data, thus enhancing the models' capability to accurately identify sensitive information while maintaining patient confidentiality. Additionally, ML models help to continuously refine the ontology [539, 540]. The models identify patterns and relationships within the data, which provide insight into potential improvements to the ontology. These insights help refine the ontology structure and contribute to a more accurate representation of the health information domain.

In essence, the interaction between ontology and ML in our work exhibits a synergistic relationship, where the strengths of one approach are tapped to complement the other [538]. The ontology model provides a meaningful and context-rich foundation for ML models, whereas the ML models contribute to the iterative refinement and validation

of the ontology. This complementary relationship culminates in a system that is not only secure and privacy-preserving, but also efficient in the management of sensitive health information [537].

5.6.3 Adoption of Privacy-Preserving Technologies for Health Information Security

Although our focus has mainly been on ontology-driven methodologies and ML techniques, it is also essential to recognize the role of privacy-preserving technologies in ensuring the security of health information. The privacy ontology model we propose in this chapter provides a structured framework for understanding and managing health-related information, but it needs to be complemented with various privacy-preserving technologies to fully realize its potential. These technologies include, but are not limited to, data encryption, differential privacy, and secure multiparty computation, which provide the technical means to protect sensitive data while still enabling valuable insights to be gleaned [541, 542]. The integration of such technologies with our proposed privacy ontology model can ensure that privacy rules and regulations, as well as the rights and privileges of stakeholders, are effectively enforced in real-world applications. This integrated approach can also address potential vulnerabilities, such as data breaches and unauthorized access, thus further enhancing the confidentiality and security of health data. Consequently, while the privacy ontology model contributes significantly to conceptualizing and organizing privacy in healthcare, the adoption of privacy-preserving technologies is integral to operationalizing these concepts and effectively safeguarding health information.

5.6.4 Section Summary

The relationship between privacy ontology and ML significantly enhances e-Healthcare systems' security, privacy, and interoperability. By offering structured and standardized frameworks, these techniques improve data management, access control, and overall system efficiency. As a result, they support the secure and confidential exchange of health information in an increasingly digital landscape.

5.7 Conclusions and Future Research

This study has explored the potential of a universal privacy model in the realm of EHR systems and context-sensitive privacy policies. Challenges such as the trade-off between

privacy and accessibility, user-friendliness, and legal compliance persist, and our work aimed to contribute to these ongoing discussions. We proposed a conceptual privacy model, employing a novel privacy ontology and an ML-based mechanism, which sought to discern between legitimate and illegitimate privacy policies while factoring in patients' PII.

We used various BERT techniques in our endeavor to pinpoint illegitimate privacy policies, indicating that our proposed ML-based approach has the potential to effectively discern such policies. Distil BERT was particularly adept at identifying both legitimate and illegitimate policies. Research suggests that refining the ML-based approach and expanding the dataset could result in a more resilient system to combat potential risks linked to inadequate privacy policies in e-Healthcare systems.

5.7.1 Limitation

Our study, while pioneering, is not without limitations. Indeed, these limitations underscore the need for further research and validation of the proposed privacy model in the context of evolving technology and privacy regulations. Our study lacks empirical evidence to fully support the effectiveness and reliability of our approach, suggesting the need for thorough evaluations and real-world testing. Moreover, the scalability and interoperability of our solution with existing healthcare IT systems and EHR platforms remain largely untested. As such, our model may require continuous updates and adjustments to align with technological advancements and emerging privacy-enhancing techniques. Ethical considerations, such as potential bias in ML algorithms, data ownership, and consent management, have yet to be explored within the scope of this chapter. Furthermore, the adaptability of our solution in the face of changing legal landscapes and ongoing compliance with changing privacy regulations warrants further scrutiny. Lastly, the practicalities of implementing our proposed solution in real-world healthcare settings, including overcoming resource constraints and resistance to change, as well as addressing the need for user training and support, are areas that require future exploration and validation.

5.7.2 Future Research Directions

To address the limitations discussed above, future research should focus on comprehensive evaluations of the proposed ontology and the ML-based approach in terms of performance, scalability, and interoperability. Future studies should also investigate strategies for integrating novel technological advancements and changes in privacy regulations to ensure the maintenance of a relevant and effective solution. Key areas of future research

to advance this field include thorough evaluations through real-world case studies and pilot implementations, exploring frameworks to adapt to advancements and changes, examining ethical implications, and fostering collaboration among stakeholders. Furthermore, future research should investigate user-centered design principles to create a solution that is user-friendly and accessible, in conjunction with the development of practical implementation strategies to seamlessly integrate the proposed solution into existing healthcare settings. The resilience of the proposed solution against various security threats and attack scenarios, along with strategies to mitigate potential vulnerabilities, remains an essential focus area for future research. By following these research directions, we hope to contribute to the ongoing evolution of secure healthcare information systems, aiming to enhance both privacy and accessibility in the world of eHealthcare.

Chapter 6

CONCLUSION

6.1 Introduction

This chapter serves to conclude the thesis by summarizing the key findings and delineating the overall influence of the research in the realm of EHR security and privacy, thus making a substantive contribution to the field. This research embarked on an extensive exploration of EHRs, with a focus on enhancing their privacy and security. The primary objective was to investigate and develop innovative techniques that combine the latest technologies and methodologies in the field. This thesis has introduced a suite of novel ideas, models, and practices, each contributing significantly to advancing the management of health data. These innovations are particularly crucial to address the multifaceted challenges currently encountered in healthcare data management, including the responsible use of Artificial Intelligence (AI) and the search for holistic solutions in healthcare data management.

In the process, this research has illuminated various aspects of EHRs, proposing new approaches and strategies to protect their privacy and security. It has shed light on the complicity in balancing accessibility and confidentiality and how emerging technologies can be leveraged to enhance this equilibrium. The thesis has also critically analyzed the role of AI and machine learning in the context of EHRs, evaluating their potential to transform healthcare data management while also considering the ethical and privacy concerns associated with their use.

The chapter also outlines the significant contributions of this research to the field. These contributions go beyond theoretical advances, providing practical insight and guidelines that healthcare professionals and policy makers can adopt. This includes a detailed examination of the challenges in implementing secure and private EHR systems, as well

as the development of frameworks and models that can be employed to overcome these challenges.

The following sections of this chapter dive into what could be improved in future research. It recognizes areas where current methodologies and technologies could be refined, suggesting a path forward for continued innovation in the security and privacy of electronic health records. This includes exploring new technologies and methodologies, as well as adapting current findings to different settings and healthcare settings.

Additionally, the chapter explores future research directions, highlighting potential areas where further investigation can yield significant advances. This includes the continuous evolution of AI and machine learning algorithms in the context of EHRs, the exploration of blockchain technology to improve data integrity and security, and the development of more robust privacy-preserving techniques. The chapter emphasizes the need for ongoing research to adapt to the rapidly changing technological landscape and evolving requirements of the healthcare industry.

In summary, this thesis provides a comprehensive overview of the current state of EHR privacy and security, offering significant contributions to the field, and laying the groundwork for future advancements. The insights and findings of this research have the potential to profoundly impact the way EHRs are managed and protected, ultimately leading to more secure and efficient healthcare systems.

6.2 Thesis Contributions

The research conducted within this thesis has made significant and multifaceted contributions to the field of EHR security and privacy. This achievement stems from a meticulous examination of carefully selected pieces of literature, strategically framed along the specific research questions that guided this study. The process of this scholarly investigation has led to the identification of several key insights and developments, which collectively enhance the understanding and advancement of EHR security and privacy. These contributions are characterized not only by their immediate relevance to the field, but also by their potential to influence future research and practice. The following sections detail these critical contributions and the research findings, emphasizing their implications in the broader context of healthcare data management.

Firstly, the research has examined the complex interaction between technological advancements and privacy concerns in EHR systems. By critically analyzing current methodologies and emerging technologies, this thesis provides a comprehensive perspective on how EHR systems can be made more secure and privacy-compliant. This includes

an exploration of the integration of novel cryptographic techniques, data anonymization processes, and the strategic use of AI and machine learning algorithms.

Secondly, the study offers a forward-looking approach to EHR security and privacy, proposing innovative models and frameworks that are poised to redefine current practices. This includes the development of a holistic security model that accommodates the evolving nature of cyber threats and the increasing sophistication of data breaches. The proposed models emphasize not only technical robustness, but also user-centric design, ensuring that EHR systems are accessible and practical for healthcare practitioners and patients alike.

Finally, the research provides a critical assessment of the policy and regulatory frameworks that govern EHR systems. By examining the interplay between technology, ethics, and law, this thesis contributes to a deeper understanding of how policy can be shaped to support the secure and ethical use of EHRs in an increasingly digital healthcare landscape.

The research in this thesis offers a series of significant contributions to the field of EHR security and privacy. These contributions, highlighted in the following, represent a blend of theoretical insight and practical application, providing a solid foundation for future advancements in the secure and responsible management of EHR.

Surveying EHRs and Privacy

A thorough in-depth analysis of EHRs, focusing on their privacy concerns, was conducted in Chapter 2. Despite their acknowledged importance, a knowledge gap was identified in the systematic understanding and management of EHRs. Existing research often interchangeably used terms such as Patient Care System (PCS), creating confusion regarding the various meanings of health records. Our survey differentiated such healthcare data related terms and addressed critical research questions. This foundational understanding set the stage for a more systematic view of EHR management, addressing data sharing methods, privacy roles, strengths and weaknesses of EHRs and technologies for privacy preservation. The identified challenges also underscore the need for future research to ensure a holistic approach to privacy, confidentiality, and security in EHRs.

Secure Data Sharing Framework EHR

A major contribution of this research was the development of a secure data sharing framework specific to the EHR system. The framework facilitates communication among healthcare stakeholders, allowing the exchange of personal and sensitive information. Through an extensive literature review, research challenges were identified and an analysis of the motivating scenario and subsequent proposal of privacy and security policies

was carried out. The chapter 3 also presented a detailed methodology, implementation plan, and future research directions to enhance the proposed framework.

Universal Privacy Model for EHR Systems

The research proposed a novel privacy model that uses privacy ontologies and machine learning techniques. Chapter 4 explored the potential of a universal privacy model for EHR systems. This model aims to balance privacy, accessibility, user-friendliness, and legitimate and illegitimate privacy policies, incorporating patient PII that address the challenges in the realm of EHR systems. This chapter significantly contributed to the ongoing discussions on context-sensitive privacy policies in EHRs.

GPT-Onto-CAABAC Framework for Advanced EHR Access Control

A significant advancement presented in this thesis (Chapter 5) is the proposed GPT-Onto-CAABAC framework. This model integrates AI, specifically GPT, with ontology and CAABAC, enhancing EHR access control and addressing the complexities of health-care data security. Despite all the challenges, the model showcases promise for diverse applications beyond healthcare, emphasizing the importance of responsible AI use.

Security Ontology Models for Health Information Systems

Chapter 2 emphasized the need for a comprehensive security ontology model for HIS. The proposed model aims to overcome challenges like scalability, user acceptance, and evolving security threats. The layered approach ensures security and scalability that offers a practical solution for HIS. Future research should focus on enhancing scalability, user-friendliness, and incorporating emerging technologies.

CEMPS Framework Implementation for EHR Privacy and Security

The chapter (Chapter 3) detailed the methodology and implementation plan for the CEMPS framework. The framework systematically identifies technologies, stakeholders, health information levels, implementing and evaluating robust security and privacy models. CEMPS aims to provide a secure and privacy-preserving environment for EHR. The chapter also highlights the various research directions for future scholars to enhance this framework further.

6.2.1 Mapping Thesis Contributions to Research Questions

This thesis has made concerted efforts to address the research questions outlined in Section 1.8, through rigorous exploration and systematic analysis. Below, we map the key contributions of this research to the respective research questions:

1. **Q1:** The integration of Access control, Blockchain, Cloud, and Cryptography technologies in enhancing EHR data sharing and access has been extensively analyzed. Chapters 2, 3, 4, and 5 demonstrate how these technologies collectively contribute to the modern healthcare landscape's efficiency and security.
2. **Q2:** The significance of privacy considerations in EHR data sharing among diverse stakeholders and its impact on the ethical and legal foundations of patient data management is critically examined. This analysis is primarily focused on Chapters 2, 3, and 4, offering insights into the correlation with distributed or centralized data management systems.
3. **Q3:** The correlation between the fundamental attributes of EHRs (Comprehensiveness, Accessibility, and Integration) and their measurable impacts in healthcare is discussed in Chapters 1, 2, and 5. This elucidates the comprehensive understanding of EHR's role in enhancing healthcare delivery.
4. **Q4:** The distinct contributions of EHR privacy, confidentiality, and security towards safeguarding patient information are quantitatively and qualitatively linked in Chapters 2, 3, 4, and 5. This establishes a robust foundation for ethical and secure healthcare data management.
5. **Q5:** The strategic integration of Access control, Blockchain, Cloud, and Cryptography (ABC) technologies to ensure the privacy and security of EHRs effectively addresses challenges related to data breaches and unauthorized access. This critical aspect is thoroughly investigated in Chapters 2, 3, 4, and 5, showcasing the optimized protection mechanisms for EHR systems.

By directly linking the contributions of this thesis to the formulated research questions, we underscore the depth and breadth of our investigation into EHR privacy and security. This mapping not only validates the comprehensive approach taken to address these complex issues but also highlights the alignment of our contributions with the core objectives of the research.

Evaluation of Thesis Contributions: The contributions of this thesis are evaluated using a diverse set of metrics, including the effectiveness of technology integration, impact on stakeholder privacy considerations, enhancement of EHR attributes, and the strengthening of data management practices. These evaluations are supported by empirical evidence and theoretical analysis, as detailed in the respective chapters, demonstrating the substantial advancements made in the field of EHR security and privacy.

6.2.2 Major Contributions to the Field

Clarification of EHR Management Concepts: This thesis significantly clarifies EHR management concepts, particularly in distinguishing between critical aspects such as privacy, confidentiality, and security (PCS). By dissecting these concepts, the research provides a subtle understanding of EHR management, highlighting the subtle, but important, differences and interconnections between these elements. This contribution helps streamline the management practices of EHR, ensuring that each aspect of PCS is adequately addressed.

Holistic Approach to EHR Security: The thesis introduces various models and frameworks that encapsulate a holistic approach to EHR security. This approach transcends traditional technical solutions, encompassing legal, ethical, and practical considerations, providing a comprehensive perspective on the security of EHRs. This multifaceted approach ensures that EHR systems are not only technically secure but also conform to legal standards and ethically sound.

Integration of Advanced Technologies: A significant stride in this research is the integration of cutting-edge technologies such as Artificial Intelligence (AI) and Machine Learning (ML) into EHR security and privacy models. This integration offers more dynamic, adaptable, and intelligent solutions to the complex challenges of EHR security, paving the way for more efficient and effective healthcare data management.

Focus on Real-world Application and Scalability: The research places a strong emphasis on developing frameworks and models that are scalable and directly applicable in real-world settings, such as the Australian MyHR system. This practical orientation ensures that the academic research conducted has tangible benefits and applicability in actual healthcare settings, enhancing the relevance and impact of the findings.

Basis for Future Research Directions: This thesis lays the foundational foundation for future research, particularly in areas such as optimizing model training, refining the integration of advanced technologies and adapting to continuously evolving regulations and threats in healthcare. It opens avenues for further exploration and innovation in EHR security and privacy, providing a robust platform on which to build on subsequent studies.

6.2.3 Implications for Practice and Policy

Enhanced EHR Security in Healthcare: The models and frameworks developed in this research provide comprehensive pathways for healthcare providers to significantly enhance

the security and privacy of EHR. These proposed solutions are not only technically robust but also align with regulatory mandates and address patient expectations for privacy and confidentiality. The detailed examination of these models contributes to a higher standard of EHR management in healthcare settings, offering strategies that are proactive and reactive to emerging security challenges.

Guidance for Implementing Advanced Security Measures: This research serves as a crucial guide for healthcare systems that intend to implement advanced security measures, particularly focusing on the integration of artificial intelligence (AI) and machine learning (ML) in the security management of the EHR. The insights provided go beyond traditional security practices, advocating for a dynamic and intelligent approach to security management that takes advantage of AI and ML capabilities. This guide is crucial for healthcare systems that navigate the complexities of modern data security in an increasingly digital healthcare landscape.

Policy Recommendations: The findings of this thesis extend to policy implications, suggesting comprehensive recommendations to standardize EHR management practices. These recommendations are particularly relevant in light of rapid technological advancements and the evolving legal and ethical landscape surrounding the access and privacy of EHRs. The thesis underscores the need for policies that are adaptable, forward-thinking, and inclusive of diverse stakeholder perspectives. Encourage policymakers to consider the complex balance between technological innovation, patient privacy, and ethical use of EHRs in formulating future healthcare policy.

6.2.4 Study Limitations

In the search for a comprehensive review of the relevant literature, this research acknowledges certain inherent limitations. A primary constraint identified was the general scarcity of literature that clearly addresses the subtle aspects of privacy preservation in EHRs, without conflating it with confidentiality and security aspects. This gap has made it challenging to draw clear distinctions and comparisons between privacy, confidentiality, and security within the context of EHR. Consequently, research encountered difficulties in pinpointing specific techniques and technologies exclusively focused on EHR privacy.

Furthermore, the validation of the validity of certain studies solely on the basis of their manuscripts presented challenges. A notable observation is the lack of empirical testing with real samples or raw EHR data in existing studies, casting doubts on their external validity. This limitation raises concerns about the applicability of the findings of these studies to real-world healthcare scenarios, as the absence of practical tests can limit the generalizability of the research conclusions.

The research gaps identified during this analysis are multifaceted. While EHR Privacy and Security are prevalent in numerous Health Systems, and emerging techniques such as Cloud computing offer promising decision-making tools based on centralized data warehousing, the extent to which these can be effectively implemented remains uncertain. The study [22] highlights that cloud analyses often encounter commonalities, making it challenging to extract representative or non-representative data without predefined privacy and security conditions. These conditions must be rigorously evaluated to determine the effectiveness of cloud computing in the management of EHR.

To improve cloud computing approaches, the recommendations and practical steps proposed in this research should be explored in conjunction with Ontology capabilities and the Semantic Web [5]. The challenges of overcoming latency in creating cross-cloud environments and providing global access to mobile users are areas that deserve improvement. The development of a general data model, with supporting policy and mapping models that link multiple data sources, is another avenue for future exploration. Privacy issues remain a significant hurdle in integrating data from various sources, with the primary goal being the balance between utility and privacy. For example, a healthcare professional might only share part of a client's data, posing a challenge in maintaining comprehensive records without privacy infringements.

An innovative approach to consider is the analysis of "similar patients" in a population, focusing on distributional aspects to maintain privacy. Although this method currently lacks precision, it has potential for future refinement and could prove useful in settings such as GP offices. Another challenge lies in the sparse and missing EHR data, which could be missing at random (MAR), missing completely at random (MCAR), or missing not at random (MNAR). Current strategies, such as omitting missing records, can lead to reduced sample sizes and potential biases, indicating the need for more sophisticated methods to handle such data discrepancies in EHR research.

6.3 Future Research Directions

The rapid advancement and increasing complexity of healthcare technology, particularly in the realm of electronic health records (EHR), require continuous research and development to address emerging challenges in security, privacy, and data sharing within healthcare data ecosystems. Building on the foundation established by this thesis, which introduced the Centralized EHR Model for Preserving Privacy and Security (CEMPS) and explored advanced technologies and ontology models, future research must pivot towards novel solutions that further these efforts. This section outlines comprehensive

directions for future research, with the aim of bridging existing gaps and fostering innovation in the management of electronic health records.

6.3.1 Advancing CEMPS with Emerging Technologies

Future studies should explore the integration of next-generation technologies with the CEMPS framework to enhance its security and privacy capabilities. This includes the application of quantum-resistant cryptography to protect against future quantum computing threats and the exploration of AI and machine learning algorithms for the detection and response of predictive threats. Additionally, the feasibility of incorporating federated learning into EHR systems for decentralized, privacy-preserving data analysis warrants thorough investigation.

6.3.2 Ontology and Machine Learning Synergies

The potential of ontology models, combined with machine learning, has been demonstrated to improve EHR security and privacy. Future research should focus on developing dynamic ontology models that can adapt to evolving healthcare terminology and privacy regulations. Furthermore, investigating machine learning algorithms' ability to automate the enforcement of complex privacy policies in real time can offer significant advancements in protecting patient data.

6.3.3 Blockchain for Decentralized Trust

Blockchain technology holds promise in creating a decentralized trust framework for EHR systems. Future work should evaluate the integration of blockchain with CEMPS to ensure tamperproof, transparent, and auditable access to health records. This includes exploring scalable consensus mechanisms suitable for healthcare data exchanges and smart contracts for automated privacy policy enforcement.

6.3.4 Ethical, Legal, and Social Implications (ELSI)

As technological solutions advance, so should our understanding of their ethical, legal, and social implications. Future research should investigate the ELSI of using advanced technologies in healthcare, focusing on patient consent mechanisms, data ownership issues, and the digital divide. The development of frameworks to guide ethical use of AI in healthcare care and the study of the legal implications of decentralized data management systems are critical.

6.3.5 Interoperability and Standardization

A significant challenge in EHR systems is ensuring interoperability among diverse health-care IT systems while maintaining high security and privacy standards. Future studies should aim at developing standardized data exchange protocols that can support the seamless and secure sharing of health information between different platforms and jurisdictions. This includes the use of open standards for data interoperability and the creation of universal APIs for EHR systems.

6.3.6 User-Centered Design and Usability

Enhancing the user experience and ensuring the widespread adoption of secure EHR systems requires a focus on user-centered design. Future research directions include conducting usability studies to identify and overcome barriers to user acceptance, designing intuitive interfaces for both healthcare providers and patients, and developing training programs to increase digital literacy in healthcare settings.

6.3.7 Comprehensive Evaluation and Real-World Implementation

Finally, there is a need for comprehensive evaluation frameworks are needed to assess the effectiveness, security, and privacy of the proposed EHR solutions in real world settings. Pilot studies and collaborations with healthcare institutions can provide valuable information on the practical challenges and benefits of implementing advanced EHR systems. In addition, longitudinal studies are essential to monitor the impact of these technologies on healthcare delivery and patient outcomes.

In conclusion, the future research directions outlined above seek to build on the contributions of this thesis toward creating more secure, efficient, and patient-centered EHR systems. By addressing these multifaceted challenges through interdisciplinary research, the field can make significant strides in improving the quality and accessibility of health-care in the digital age.

Bibliography

- [1] Kamrun Nahar and Asif Qumer Gill. Integrated identity and access management metamodel and pattern system for secure enterprise architecture. *Data & Knowledge Engineering*, 140:102038–102038, 2022.
- [2] Mary Brown and Karen White. Securing electronic health records: A critical analysis of current measures and future directions. *Journal of Medical Privacy*, 19(1):12–29, 2023.
- [3] Qinyong Lin, Xiaorong Li, Ken Cai, Mohan Prakash, and D Paulraj. Secure internet of medical things (iomt) based on ecmqv-mac authentication protocol and ekmc-scp blockchain networking. *Information Sciences*, 654:119783, 2024.
- [4] M Lakshmanan, GS Anandha Mala, and KM Anandkumar. Highly secured ehr management system based on blockchain technology with digitally signed authentication using data sanitization and polynomial interpolation. *Biomedical Signal Processing and Control*, 87:105412, 2024.
- [5] Dewei Yang, Shuai Liu, Jinbao Sheng, Xuehui Peng, and Huiwen Wang. Construction and implementation of ontology model of dam break emergency plan. *Advances in Measurement Technology, Disaster Prevention and Mitigation*, pages 213–220, 2023.
- [6] José Luis Fernández-Alemán, Inmaculada Carrión Señor, Pedro Ángel Oliver Lozoya, and Ambrosio Toval. Security and privacy in electronic health records: A systematic literature review. *Journal of biomedical informatics*, 46(3):541–562, 2013.
- [7] Clemens Scott Kruse, Benjamin Frederick, Taylor Jacobson, and D Kyle Monticone. Cybersecurity in healthcare: A systematic review of modern threats and trends. *Technology and Health Care*, 25(1):1–10, 2017.
- [8] Ehab Zaghoul, Tongtong Li, and Jian 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)*, pages 375–379. IEEE, 2019.
- [9] Asad Ali Siyal, Aisha Zahid Junejo, Muhammad Zawish, Kainat Ahmed, Aiman Khalil, and Georgia Soursou. Applications of blockchain technology in medicine and healthcare: Challenges and future perspectives. *Cryptography*, 3(1):3, 2019.
- [10] Hao Wang and Yujiao Song. Secure cloud-based ehr system using attribute-based cryptosystem and blockchain. *Journal of medical systems*, 42(8):1–9, 2018.
- [11] Gaby G Dagher, Jordan Mohler, Matea Milojkovic, and Praneeth Babu Marella. Ancile: Privacy-preserving framework for access control and interoperability of electronic health records using blockchain technology. *Sustainable cities and society*, 39:283–297, 2018.
- [12] Mohammad Shahid Husain, Muhamad Hariz Bin Muhamad Adnan, Mohammad Zunnun Khan, Saurabh Shukla, and Fahad U Khan. *Pervasive Healthcare: A Compendium of Critical Factors for Success*. Springer, 2021.
- [13] Hsuan-Yu Chen, Zhen-Yu Wu, Tzer-Long Chen, Yao-Min Huang, and Chia-Hui Liu. Security privacy and policy for cryptographic based electronic medical information system. *Sensors*, 21(3):713, 2021.
- [14] Muhammad Anshari. Redefining electronic health records (ehr) and electronic medical records (emr) to promote patient empowerment. *IJID (International Journal on Informatics for Development)*, 8(1):35–39, 2019.
- [15] Maithilee Joshi, Karuna Joshi, and Tim Finin. Attribute based encryption for secure access to cloud based ehr systems. In *2018 IEEE 11th International Conference on Cloud Computing (CLOUD)*, pages 932–935. IEEE, 2018.
- [16] Perihan Elif Ekmekci and Berna Arda. Interculturalism and informed consent: Respecting cultural differences without breaching human rights. *Cultura*, 14(2):159–172, 2017.
- [17] Anastassia Negrouk, Denis Horgan, Alessandra Gorini, Ilaria Cutica, Lada Leyens, Sebastian Schee genannt Halfmann, and Gabriella Pravettoni. Clinical trials, data protection and patient empowerment in the era of the new eu regulations. *Public health genomics*, 18(6):386–395, 2015.
- [18] D Townend. *Implementation of the Data Protection Directive in relation to medical research in Europe*, 2017. Routledge.
- [19] E Bender. Big data in biomedicine: 4 big questions. *Nature*, 527(7576):19–19, 2015.

- [20] Xinghua Shi and Xintao Wu. An overview of human genetic privacy. *Annals of the New York Academy of Sciences*, 1387(1):61–72, 2017.
- [21] Robert B Franke, Paul A Carson, and Brett J Enyeart. Ehr confidentiality, integrity, and availability: Hipaa and hitech compliance. In *2018 IEEE International Symposium on Technologies for Homeland Security (HST)*, pages 1–7. IEEE, 2018.
- [22] Author(s) Name. Title of the differential privacy in healthcare paper. *Journal Name*, 2021.
- [23] Ketan Paranjape, Michiel Schinkel, and Prabath Nanayakkara. Short keynote paper: Mainstreaming personalized healthcare—transforming healthcare through new era of artificial intelligence. *IEEE journal of biomedical and health informatics*, 24(7):1860–1863, 2020.
- [24] Sasan Ghorbani Kalkhajeh, Azam Aghajari, Behnaz Dindamal, Zohreh Shahvali-Kuhshuri, and Farzad Faraji-Khiavi. The integrated electronic health system in iranian health centers: benefits and challenges. *BMC Primary Care*, 24(1):53, 2023.
- [25] Rui Zhang and Ling Liu. Security models and requirements for healthcare application clouds. In *Proceedings of the 2010 IEEE 3rd International Conference on Cloud Computing*, pages 268–275. IEEE, 2010.
- [26] Sunday Adeola Ajagbe, AO Adesina, and JB Oladosu. Empirical evaluation of efficient asymmetric encryption algorithms for the protection of electronic medical records (emr) on web application. *International Journal of Scientific and Engineering Research*, 10(5):848–871, 2019.
- [27] J Randall Curtis, Seelwan Sathitratanacheewin, Helene Starks, Robert Y Lee, Erin K Kross, Lois Downey, James Sibley, William Lober, Elizabeth T Loggers, James A Fausto, et al. Using electronic health records for quality measurement and accountability in care of the seriously ill: opportunities and challenges. *Journal of palliative medicine*, 21(S2):S–52, 2018.
- [28] Joan M Teno, Rebecca Anhang Price, and Lena K Makaroun. Challenges of measuring quality of community-based programs for seriously ill individuals and their families. *Health Aff (Millwood)*, 36(7):1227–1233, 2017. doi: 10.1377/hlthaff.2017.0161.
- [29] David Blumenthal and Melinda K Abrams. Tailoring complex care management for high-need, high-cost patients. *JAMA*, 316(16):1657–1658, 2016. doi: 10.1001/jama.2016.12388.

- [30] Julie P. W. Bynum, Andrea Austin, Donald Carmichael, and Ellen Meara. High-cost dual-eligibles' service use demonstrates need for supportive and palliative models of care. *Health Aff (Millwood)*, 36(7):1309–1317, 2017. doi: 10.1377/hlthaff.2017.0157.
- [31] Melissa D Aldridge and Elizabeth H Bradley. Epidemiology and patterns of care at the end of life: Rising complexity, shifts in care patterns and sites of death. *Health Aff (Millwood)*, 36(7):1175–1183, 2017. doi: 10.1377/hlthaff.2017.0182.
- [32] Suchitra Kataria and Vinod Ravindran. Electronic health records: a critical appraisal of strengths and limitations. *Journal of the Royal College of Physicians of Edinburgh*, 50(3):262–268, 2020.
- [33] Yi Liu, Yinghui Zhang, Jie Ling, and Zhusong Liu. Office of the national coordinator for health information technology (onc), 2018. URL <https://www.healthit.gov/sites/default/files/utility/final-federal-health-it-strategic-plan-0911.pdf>. Federal Health Information Technology Strategic Plan.
- [34] Sharon L. Adams. Nurses knowledge, skills, and attitude toward electronic health records (ehr). *Walden University*, 2015.
- [35] Nir Menachemi and Taleah H Collum. Benefits and drawbacks of electronic health record systems. *Risk Manag Healthc Policy*, 4:47–55, 2011.
- [36] Ilias Maglogiannis. Towards the adoption of open source and open access electronic health record systems. *Journal of Healthcare Engineering*, 3:141–161, 2012. doi: <https://doi.org/10.1260/2040-2295.3.1.141>.
- [37] Hon. Tony Blair. The challenge for the nhs is to harness the information revolution and use it to benefit patients, 2005. URL https://webarchive.nationalarchives.gov.uk/ukgwa/20120503231618/http://www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/@dh/@en/documents/digitalasset/dh_4014469.pdf.
- [38] Bob Gann. Information for health. *Health Expect*, 2(1):72, 1999. doi: 10.1046/j.1369-6513.1999.0032c.x.
- [39] Jessica S Ancker, Lisa M Kern, Alison Edwards, Sarah Nosal, Daniel M Stein, Diane Hauser, Rainu Kaushal, and HITEC Investigators. Associations between healthcare quality and use of electronic health record functions in ambulatory care. *J Am Med Inform Assoc*, 22(4):864–71, 2015.
- [40] Australian Digital Health Agency, 2022. URL www.conversation.digitalhealth.gov.au. Australia's National Digital Health Strategy.

- [41] Leila Shahmoradi, Alireza Darrudi, Goli Arji, and Ahmadreza Farzaneh Nejad. Electronic health record implementation: a swot analysis. *Acta Medica Iranica*, pages 642–649, 2017.
- [42] Xiaoxun Sun, Hua Wang, Jiuyong Li, and Jian Pei. Publishing anonymous survey rating data. *Data Mining and Knowledge Discovery*, 23:379–406, 11 2011. doi: 10.1007/s10618-010-0208-4.
- [43] Le Sun, Jiangan Ma, Hua Wang, and Yanchun Zhang. Cloud service description model: An extension of usdl for cloud services. *IEEE Transactions on Services Computing*, PP:1–1, 08 2015. doi: 10.1109/TSC.2015.2474386.
- [44] Leila Shahmoradi, Alireza Darrudi, Goli Arji, and Ahmadreza Farzaneh Nejad. Electronic health record implementation: A swot analysis. *Acta Med Iran*, 55(10): 642–649, 2017.
- [45] Liang Chen, Martin J Kollingbaum, Timothy J Norman, and Peter Edwards. Risk-aware access control for electronic health records. In *Proceedings of the Third Annual Digital Economy All Hands Conference, Aberdeen*, 2012.
- [46] HIPPA. Health information privacy. <https://www.hhs.gov/hipaa/index.html>, 2021. Accessed: 2021-10-01.
- [47] Peter M Sfikas. Hipaa security regulations: protecting patients’ electronic health information. *The Journal of the American Dental Association*, 134(5):640–643, 2003.
- [48] Raza Nowrozy, Khandakar Ahmed, ASM Kayes, Hua Wang, and Timothy R McIntosh. Privacy preservation of electronic health records in the modern era: A systematic survey. *ACM Computing Surveys*, 2024.
- [49] Yong-Feng Ge, Hua Wang, Elisa Bertino, Zhi-Hui Zhan, Jinli Cao, Yanchun Zhang, and Jun Zhang. Evolutionary dynamic database partitioning optimization for privacy and utility. *IEEE Transactions on Dependable and Secure Computing*, 2023.
- [50] Siuly Siuly, Ömer Faruk Alçın, Hua Wang, Yan Li, and Peng Wen. Exploring rhythms and channels-based eeg biomarkers for early detection of alzheimer’s disease. *IEEE Transactions on Emerging Topics in Computational Intelligence*, 2024.
- [51] Jingshan Li and Pascale Carayon. Health care 4.0: A vision for smart and connected health care. *IISE Transactions on Healthcare Systems Engineering*, 11(3):171–180, 2021.
- [52] HealthIT.gov. Technical safeguards. <https://www.healthit.gov/faq/what-are-technical-safeguards>, 2021.

- [53] Nduma N Basil, Solomon Ambe, Chukwuyem Ekhatior, Ekokobe Fonkem, Basil N Nduma, and Chukwuyem Ekhatior. Health records database and inherent security concerns: A review of the literature. *Cureus*, 14(10), 2022.
- [54] Mike Chapple and David Seidl. *Cyberwarfare: Information operations in a connected world*. Jones & Bartlett Learning, 2021.
- [55] Waseem Ahmed Khattak and Fazle Rabbi. Ethical considerations and challenges in the deployment of natural language processing systems in healthcare. *International Journal of Applied Health Care Analytics*, 8(5):17–36, 2023.
- [56] Arianna Dagliati, Alberto Malovini, Valentina Tibollo, and Riccardo Bellazzi. Health informatics and ehr to support clinical research in the covid-19 pandemic: an overview. *Briefings in bioinformatics*, 22(2):812–822, 2021.
- [57] Hyowon Im, Ki-Hyung Kim, and Jai-Hoon Kim. Privacy and ledger size analysis for healthcare blockchain. In *2020 International Conference on Information Networking (ICOIN)*, pages 825–829. IEEE, 2020.
- [58] Ann Cavoukian. Privacy by design in the age of big data analytics and ai. *Journal of the American Medical Informatics Association*, 27(11):1756–1759, 2020.
- [59] Fei Zhu, Xun Yi, Alsharif Abuadbba, Ibrahim Khalil, Xinyi Huang, and Feihong Xu. A security-enhanced certificateless conditional privacy-preserving authentication scheme for vehicular ad hoc networks. *IEEE Transactions on Intelligent Transportation Systems*, 2023.
- [60] David Moher, Alessandro Liberati, Jennifer Tetzlaff, and Douglas G. Altman. Preferred reporting items for systematic reviews and meta-analyses: The prisma statement. *International Journal of Surgery*, 8(5):336–341, 2020.
- [61] Md Nurul Ahad Tawhid, Siuly Siuly, Kate Wang, and Hua Wang. Automatic and efficient framework for identifying multiple neurological disorders from eeg signals. *IEEE Transactions on Technology and Society*, 4(1):76–86, 2023.
- [62] Chuan Wang, Bing Sun, Ke-Jing Du, Jian-Yu Li, Zhi-Hui Zhan, Sang-Woon Jeon, Hua Wang, and Jun Zhang. A novel evolutionary algorithm with column and sub-block local search for sudoku puzzles. *IEEE Transactions on Games*, 2023.
- [63] Jia-Quan Yang, Qi-Te Yang, Ke-Jing Du, Chun-Hua Chen, Hua Wang, Sang-Woon Jeon, Jun Zhang, and Zhi-Hui Zhan. Bi-directional feature fixation-based particle swarm optimization for large-scale feature selection. *IEEE Transactions on Big Data*, 2022.

- [64] Jiao Yin, MingJian Tang, Jinli Cao, Mingshan You, Hua Wang, and Mamoun Alazab. Knowledge-driven cybersecurity intelligence: Software vulnerability coexploitation behavior discovery. *IEEE transactions on industrial informatics*, 19(4): 5593–5601, 2022.
- [65] Ashik Mostafa Alvi, Siuly Siuly, and Hua Wang. A long short-term memory based framework for early detection of mild cognitive impairment from eeg signals. *IEEE Transactions on Emerging Topics in Computational Intelligence*, 7(2): 375–388, 2022.
- [66] Jian-Yu Li, Ke-Jing Du, Zhi-Hui Zhan, Hua Wang, and Jun Zhang. Distributed differential evolution with adaptive resource allocation. *IEEE transactions on cybernetics*, 2022.
- [67] Wen Shi, Wei-Neng Chen, Sam Kwong, Jie Zhang, Hua Wang, Tianlong Gu, Huaqiang Yuan, and Jun Zhang. A coevolutionary estimation of distribution algorithm for group insurance portfolio. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 52(11):6714–6728, 2021.
- [68] Yong-Feng Ge, Wei-Jie Yu, Jinli Cao, Hua Wang, Zhi-Hui Zhan, Yanchun Zhang, and Jun Zhang. Distributed memetic algorithm for outsourced database fragmentation. *IEEE Transactions on Cybernetics*, 51(10):4808–4821, 2020.
- [69] Siuly Siuly, Smith K Khare, Varun Bajaj, Hua Wang, and Yanchun Zhang. A computerized method for automatic detection of schizophrenia using eeg signals. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 28(11):2390–2400, 2020.
- [70] Jian-Yu Li, Zhi-Hui Zhan, Hua Wang, and Jun Zhang. Data-driven evolutionary algorithm with perturbation-based ensemble surrogates. *IEEE Transactions on Cybernetics*, 51(8):3925–3937, 2020.
- [71] GDPR. The european general data protection regulation. <https://gdpr.eu/what-is-gdpr>, 2018. Accessed: 2021-10-01.
- [72] Sandra Wachter. The gdpr and the internet of things: a three-step transparency model. *Law, Innovation and Technology*, 10(2):266–294, 2018.
- [73] Aikaterini Daoultzoglou. Gdpr in education. 2022.
- [74] Konrad Magnus. The importance of the cia triad in the age of the internet. *MATEC Web of Conferences*, 329:03035, 2020.
- [75] Office for Civil Rights. Administrative safeguards. <https://www.hhs.gov/hipaa/for-professionals/security/administrative-safeguards/index.html>, 2021.

- [76] Office for Civil Rights. Individual rights. <https://www.hhs.gov/hipaa/for-individuals/index.html>, 2021.
- [77] P Kostkova, H Brewer, S De Lusignan, E Fottrell, B Goldacre, G Hart, P Koczan, P Knight, C Marsolier, RA McKendry, et al. Who owns the data? open data for healthcare. *front. Public Health*, 4(7):1–6, 2016.
- [78] Christian Esposito, Alfredo De Santis, Genny Tortora, Henry Chang, and Kim-Kwang Raymond Choo. Blockchain: A panacea for healthcare cloud-based data security and privacy? *IEEE Cloud Computing*, 5(1):31–37, 2018.
- [79] Katherine K Kim, Jill G Joseph, and Lucila Ohno-Machado. Comparison of consumers’ views on electronic data sharing for healthcare and research. *Journal of the American Medical Informatics Association*, 22(4):821–830, 2015.
- [80] Hal Berghel. Equifax and the latest round of identity theft roulette. *Computer*, 50(12):72–76, 2017.
- [81] Shi Dong, Khushnood Abbas, and Raj Jain. A survey on distributed denial of service (ddos) attacks in sdn and cloud computing environments. *IEEE Access*, 7: 80813–80828, 2019.
- [82] Mayra Rosario Fuentes. Cybercrime and other threats faced by the healthcare industry. *Trend Micro*, 5566, 2017.
- [83] Eman AbuKhoua, Nader Mohamed, and Jameela Al-Jaroodi. e-health cloud: opportunities and challenges. *Future internet*, 4(3):621–645, 2012.
- [84] Sagar Sharma, Keke Chen, and Amit Sheth. Toward practical privacy-preserving analytics for iot and cloud-based healthcare systems. *IEEE Internet Computing*, 22(2):42–51, 2018.
- [85] Benjamin Fabian, Tatiana Ermakova, and Philipp Junghanns. Collaborative and secure sharing of healthcare data in multi-clouds. *Information Systems*, 48:132–150, 2015.
- [86] William J Gordon and Christian Catalini. Blockchain technology for healthcare: facilitating the transition to patient-driven interoperability. *Computational and structural biotechnology journal*, 16:224–230, 2018.
- [87] Michelle Finneran Denedy, Jonathan Fox, and Thomas R Finneran. Data and privacy governance concepts. In *The Privacy Engineer’s Manifesto*, pages 51–72. Springer, 2014.

- [88] John Smith and Linda Johnson. The impact of ehr breaches on healthcare providers and patients. *Journal of Cybersecurity*, 17(3):45–61, 2022.
- [89] Rashaad ET Jones, Erik S Connors, Mary E Mossey, John R Hyatt, Neil J Hansen, and Mica R Endsley. Using fuzzy cognitive mapping techniques to model situation awareness for army infantry platoon leaders. *Computational and Mathematical Organization Theory*, 17(3):272–295, 2011.
- [90] Alice Doe and Robert Smith. Vulnerabilities in ehr systems: A growing concern in the digital age. *Health Informatics Review*, 13(2):22–37, 2021.
- [91] DonHee Lee and Seong No Yoon. Application of artificial intelligence-based technologies in the healthcare industry: Opportunities and challenges. *International Journal of Environmental Research and Public Health*, 18(1):271, 2021.
- [92] Erik Rissanen and Rui Yang. Analyzing the complexity and performance of abac policies. *IEEE Transactions on Dependable and Secure Computing*, 16(2), 2019.
- [93] Dhruv Pandit and R Sudhakar. Enhanced security and privacy of healthcare big data using blockchain technology: A review. *Journal of medical systems*, 45(4):36, 2021.
- [94] Xueyu Geng, Yuhong Qian, Jing Zhang, Qiong Jiang, and Lijuan Wang. Intelligent medical data sharing and privacy protection based on blockchain and ipfs. *Journal of medical systems*, 43(5):107, 2019.
- [95] Sofia Terzi and Ioannis Stamelos. Architectural solutions for improving transparency, data quality, and security in ehealth systems by designing and adding blockchain modules, while maintaining interoperability: the ehdsi network case. *Health and Technology*, pages 1–12, 2024.
- [96] Mujeeb C Kandy, Jazeel Abdulmajeed, Chiragkumar N Gohel, John M Gibb, and Mohamed Ghaith Al-Kuwari. Transforming primary healthcare services with centralized health intelligence: A case study from qatar. *Qatar Journal of Public Health*, 2023(2):8, 2024.
- [97] Sumana Srivatsa, David Walsh, Mayada Aljehani, Jason Weinreb, Nathan Becker, Benjamin H Ellis, Renee George, Xingyao Chen, Naim Matasci, Reva Basho, et al. Causal personalized treatment estimation framework using real-world electronic health record (ehr) data to inform cancer care decisions. *Cancer Research*, 84 (6_Supplement):6467–6467, 2024.

- [98] Seun Solomon Bakare, Adekunle Oyeyemi Adeniyi, Chidiogo Uzoamaka Akpuokwe, and Nkechi Emmanuella Eneh. Data privacy laws and compliance: A comparative review of the eu gdpr and usa regulations. *Computer Science & IT Research Journal*, 5(3):528–543, 2024.
- [99] Joshua D Symons, Hutan Ashrafian, Rachel Dunscombe, and Ara Darzi. From ehr to phr: let’s get the record straight. *BMJ open*, 9(9):e029582, 2019.
- [100] Ismai IKeshta and Ammar Odeh. Security and privacy of electronic health records: Concerns and challenges. *Egyptian Informatics Journal*, 22(2):117–183, 2021.
- [101] Alevtina Dubovitskaya, Zhigang Xu, Samuel Ryu, Michael Schumacher, and Fusheng Wang. Secure and trustable electronic medical records sharing using blockchain. In *AMIA annual symposium proceedings*, volume 2017, page 650. American Medical Informatics Association, 2017.
- [102] D Akarca, PY Xiu, D Ebbitt, B Mustafa, H Al-Ramadhani, and A Albeyatti. Blockchain secured electronic health records: Patient rights, privacy and cybersecurity. In *2019 10th International Conference on Dependable Systems, Services and Technologies (DESSERT)*, pages 108–111. IEEE, 2019.
- [103] Aaron Adler, Michael J Mayhew, Jeffrey Cleveland, Michael Atighetchi, and Rachel Greenstadt. Using machine learning for behavior-based access control: Scalable anomaly detection on tcp connections and http requests. In *MILCOM 2013-2013 IEEE Military Communications Conference*, pages 1880–1887. IEEE, 2013.
- [104] Fatemeh Rezaeibagha, Yi Mu, Willy Susilo, and Khin Than Win. Multi-authority security framework for scalable ehr systems. *International Journal of Medical Engineering and Informatics*, 8(4):390–408, 2016.
- [105] Nir Kshetri. Blockchain’s roles in strengthening cybersecurity and protecting privacy. *Telecommunications policy*, 41(10):1027–1038, 2017.
- [106] Pasupathy Vimalachandran, Yanchun Zhang, Jinli Cao, Lili Sun, and Jianming Yong. Preserving data privacy and security in australian my health record system: A quality health care implication. In *International Conference on Web Information Systems Engineering*, pages 111–120. Springer, 2018.
- [107] A H Omar. The Effect of Electronic Health Records on Undergraduate and Postgraduate Medical Education: A Scoping Review. *Canada*, 2019.
- [108] N Kshetri. Blockchain’s roles in strengthening cybersecurity and protecting privacy. *Telecommunication Policy*, 41(10):1027–1038, 2017.

- [109] E Zaghoul, T Li, and J Ren. Security and Privacy of Electronic Health Records: Decentralized and Hierarchical Data Sharing using Smart Contracts. *2019 International Conference on Computing, Networking and Communications (ICNC)*, pages 375–379, 2019.
- [110] A S Kayes, W Rahayu, T Dillon, E Chang, and J Han. Context-aware access control with imprecise context characterization for cloud-based data resources. *Future Generation Computer Systems*, 93:237–255, 2019.
- [111] K Munir, M Odeh, and R McClatchey. Ontology-driven relational query formulation using the semantic and assertional capabilities of OWL-DL. *Knowledge- Based Systems*, 35:144–159, 2012.
- [112] M Van Der Haak, A C Wolff, R Brandner, P Drings, M Wannemacher, and T Wetter. Data security and protection in cross-institutional electronic patient records. *International journal of medical informatics*, 70(2-3):117–130, 2003.
- [113] Abdullah Al Omar, Md Zakirul Alam Bhuiyan, Anirban Basu, Shinsaku Kiyomoto, and Mohammad Shahriar Rahman. Privacy-friendly platform for healthcare data in cloud based on blockchain environment. *Future generation computer systems*, 95:511–521, 2019.
- [114] Craig Fink. Privacy and confidentiality in the virtual classroom: instructor perceptions, knowledge and strategies. 2012.
- [115] Yong Wang, Aiqing Zhang, Peiyun Zhang, and Huaqun Wang. Cloud-assisted ehr sharing with security and privacy preservation via consortium blockchain. *Ieee Access*, 7:136704–136719, 2019.
- [116] Garima Verma. Blockchain-based privacy preservation framework for healthcare data in cloud environment. *Journal of Experimental & Theoretical Artificial Intelligence*, 36(1):147–160, 2024.
- [117] André Henrique Mayer, Cristiano André da Costa, and Rodrigo da Rosa Righi. Electronic health records in a blockchain: A systematic review. *Health informatics journal*, 26(2):1273–1288, 2020.
- [118] Hua Wang, Yanchun Zhang, and Jinli Cao. Ubiquitous computing environments and its usage access control. volume 152, page 6, 01 2006. doi: 10.1145/1146847.1146853.
- [119] Jiao Yin, Mingjian Tang, Jinli Cao, Hua Wang, Mingshan You, and Yongzheng Lin. Vulnerability exploitation time prediction: an integrated framework for dynamic

- imbalanced learning. *World Wide Web*, pages 401–423, 01 2022. doi: 10.1007/s11280-021-00909-z.
- [120] Jagmeet Singh Aidan, Harsh Kumar Verma, and Lalit Kumar Awasthi. Comprehensive survey on petya ransomware attack. In *2017 International Conference on Next Generation Computing and Information Systems (ICNGCIS)*, pages 122–125. IEEE, 2017.
- [121] Timothy McIntosh, ASM Kayes, Yi-Ping Phoebe Chen, Alex Ng, and Paul Waters. Ransomware mitigation in the modern era: A comprehensive review, research challenges, and future directions. *ACM Computing Surveys (CSUR)*, 54(9):1–36, 2021.
- [122] Sudeep Tanwar, Karan Parekh, and Richard Evans. Blockchain-based electronic healthcare record system for healthcare 4.0 applications. *Journal of Information Security and Applications*, 50:102407, 2020.
- [123] Maryam Farhadi, Hisham Haddad, and Hossain Shahriar. Static analysis of hipaa security requirements in electronic health record applications. In *2018 IEEE 42nd Annual Computer Software and Applications Conference (COMPSAC)*, volume 2, pages 474–479. IEEE, 2018.
- [124] Barbara Kitchenham. Procedures for performing systematic reviews. *Keele, UK, Keele University*, 33(2004):1–26, 2004.
- [125] Raza Nowrozy, Ahmed Khandakar, Wang Hua, and Timothy Mcintosh. Towards a universal privacy model for electronic health record systems: An ontology and machine learning approach. *Informatics*, 10(3), 2023.
- [126] Vivek Subbiah. The next generation of evidence-based medicine. *Nature medicine*, 29(1):49–58, 2023.
- [127] Roberto Cerchione, Piera Centobelli, Emanuela Riccio, Stefano Abbate, and Eugenio Oropallo. Blockchain’s coming to hospital to digitalize healthcare services: Designing a distributed electronic health record ecosystem. *Technovation*, 120: 102480, 2023.
- [128] Bassim Al Bahrani, Itrat Medhi, and ITRAT MEHDI. Copy-pasting in patients’ electronic medical records (emrs): Use judiciously and with caution. *Cureus*, 15(6), 2023.
- [129] Nilüfer Demirsoy and Nurdan Kirimlioglu. Protection of privacy and confidentiality as a patient right: physicians’ and nurses’ viewpoints. *Biomedical Research*, 27(4): 1437–1448, 2016.

- [130] John D Halamka, Andrew Lippman, and Ariel Ekblaw. The potential for blockchain to transform electronic health records. *Harvard Business Review*, 3(3):2–5, 2017.
- [131] Tasha Glenn and Scott Monteith. Privacy in the digital world: medical and health data outside of hipaa protections. *Current psychiatry reports*, 16(11):1–11, 2014.
- [132] Aisling R Caffrey and Austin R Horn. Considerations for protecting research participants. In *Pragmatic Randomized Clinical Trials*, pages 273–292. Elsevier, 2021.
- [133] Amanda M Gutierrez, Jacob D Hofstetter, Emma L Dishner, Elizabeth Chiao, Dilreet Rai, and Amy L McGuire. A right to privacy and confidentiality: ethical medical care for patients in united states immigration detention. *Journal of Law, Medicine & Ethics*, 48(1):161–168, 2020.
- [134] W Bani Issa, I Al Akour, A Ibrahim, A Almarzouqi, S Abbas, F Hisham, and J Griffiths. Privacy, confidentiality, security and patient safety concerns about electronic health records. *International Nursing Review*, 67(2):218–230, 2020.
- [135] Gabriela Kato Lettieri, Aline Hung Tai, Aline Rodrigues Hütter, André Luiz Torres Raszl, Mariana Moura, and Raquel Barbosa Cintra. Medical confidentiality in the digital era: an analysis of physician-patient relations. *Revista Bioetica*, 29:814–824, 2022.
- [136] BD Deebak and Fadi Al-Turjman. Secure-user sign-in authentication for iot-based ehealth systems. *Complex & Intelligent Systems*, 9(3):2629–2649, 2023.
- [137] Raza Nowrozy and Khandakar Ahmed. Enhancing health information systems security: An ontology model approach. In *International Conference on Health Information Science*, pages 91–100. Springer, 2023.
- [138] Nancy D Harada, Laural Traylor, Kathryn Wirtz Rugen, Judith L Bowen, C Scott Smith, Bradford Felker, Deborah Ludke, Ivy Tonnu-Mihara, Joshua L Ruberg, Jayson Adler, et al. Interprofessional transformation of clinical education: the first six years of the veterans affairs centers of excellence in primary care education. *Journal of interprofessional care*, 37(sup1):S86–S94, 2023.
- [139] Fatemeh Rezaeibagha, Khin Than Win, and Willy Susilo. A systematic literature review on security and privacy of electronic health record systems: technical perspectives. *Health Information Management Journal*, 44(3):23–38, 2015.
- [140] Sreelakshmi Krishnamoorthy, Amit Dua, and Shashank Gupta. Role of emerging technologies in future iot-driven healthcare 4.0 technologies: A survey, current challenges and future directions. *Journal of Ambient Intelligence and Humanized Computing*, 14(1):361–407, 2023.

- [141] Sarah Qahtan, Khaironi Yatim, Hazura Zulzalil, Mohd Hafeez Osman, AA Zaidan, and HA Alsattar. Review of healthcare industry 4.0 application-based blockchain in terms of security and privacy development attributes: Comprehensive taxonomy, open issues and challenges and recommended solution. *Journal of Network and Computer Applications*, 209:103529, 2023.
- [142] Gurvirender PS Tejay and Zareef A Mohammed. Cultivating security culture for information security success: A mixed-methods study based on anthropological perspective. *Information & Management*, 60(3):103751, 2023.
- [143] Xiaoshuai Zhang and Stefan Poslad. Blockchain support for flexible queries with granular access control to electronic medical records (emr). In *2018 IEEE International conference on communications (ICC)*, pages 1–6. IEEE, 2018.
- [144] Mohammad Mohammadi, Bagher Larijani, Seyed Hassan Emami Razavi, Akbar Fotouhi, Ahmad Ghaderi, Seyed Javad Madani, and Mohammad Naser Shafiee. Do patients know that physicians should be confidential? study on patients' awareness of privacy and confidentiality. *Journal of medical ethics and history of medicine*, 11, 2018.
- [145] Xu Yang, Xuechao Yang, Junwei Luo, Xun Yi, Ibrahim Kahlil, Shangqi Lai, Wei Wu, and Albert Y Zomaya. Towards sustainable trust: A practical sgx aided anonymous reputation system. *IEEE Transactions on Sustainable Computing*, 2023.
- [146] Sandeep Kumar Panda, Ajay Kumar Jena, Santosh Kumar Swain, and Suresh Chandra Satapathy. *Blockchain Technology: Applications and Challenges*. Springer, 2021.
- [147] Thein Than Thwin and Sangsuree Vasupongayya. Blockchain based secret-data sharing model for personal health record system. In *2018 5th International Conference on Advanced Informatics: Concept Theory and Applications (ICAICTA)*, pages 196–201. IEEE, 2018.
- [148] Rui Zhang, Rui Xue, and Ling Liu. Security and privacy on blockchain. *ACM Computing Surveys (CSUR)*, 52(3):1–34, 2019.
- [149] Yongsan Ma, Gang Zhou, and Shuangquan Wang. Wifi sensing with channel state information: A survey. *ACM Computing Surveys (CSUR)*, 52(3):1–36, 2019.
- [150] Matthew Pirretti, Patrick Traynor, Patrick McDaniel, and Brent Waters. Secure attribute-based systems. In *Proceedings of the 13th ACM conference on Computer and communications security*, pages 99–112, 2006.

- [151] Imran Makhdoom, Mehran Abolhasan, Haider Abbas, and Wei Ni. Blockchain's adoption in iot: The challenges, and a way forward. *Journal of Network and Computer Applications*, 125:251–279, 2019.
- [152] Tareq Ahram, Arman Sargolzaei, Saman Sargolzaei, Jeff Daniels, and Ben Amaba. Blockchain technology innovations. In *2017 IEEE Technology & Engineering Management Conference (TEMSCON)*, pages 137–141. IEEE, 2017.
- [153] Md Mehedi Hassan Onik, Satyabrata Aich, Jinhong Yang, Chul-Soo Kim, and Hee-Cheol Kim. Blockchain in healthcare: Challenges and solutions. In *Big data analytics for intelligent healthcare management*, pages 197–226. Elsevier, 2019.
- [154] Marek A Cyran. Blockchain as a foundation for sharing healthcare data. *Blockchain in Healthcare Today*, 2018.
- [155] Lanxiang Chen, Wai-Kong Lee, Chin-Chen Chang, Kim-Kwang Raymond Choo, and Nan Zhang. Blockchain based searchable encryption for electronic health record sharing. *Future generation computer systems*, 95:420–429, 2019.
- [156] Clemens Scott Kruse, Benjamin Smith, Hannah Vanderlinden, and Alexandra Nealand. Security techniques for the electronic health records. *Journal of Medical Systems*, 41(8):127, 2017.
- [157] Thore Graepel, Kristin Lauter, and Michael Naehrig. Ml confidential: Machine learning on encrypted data. In *International Conference on Information Security and Cryptology*, pages 1–21. Springer, 2012.
- [158] Protection Regulation. Regulation (eu) 2016/679 of the european parliament and of the council. *Regulation (eu)*, 679:2016, 2016.
- [159] Rakesh Agrawal, Jerry Kiernan, Ramakrishnan Srikant, and Yirong Xu. Hippocratic databases. In *VLDB'02: Proceedings of the 28th International Conference on Very Large Databases*, pages 143–154. Elsevier, 2002.
- [160] Daisuke Mashima. *Safeguarding health data with enhanced accountability and patient awareness*. Georgia Institute of Technology, 2012.
- [161] Chang Sun, Lianne Ippel, Johan Van Soest, Birgit Wouters, Alexander Malic, Onaopepo Adekunle, Bob van den Berg, Ole Mussmann, Annemarie Koster, Carla van der Kallen, et al. A privacy-preserving infrastructure for analyzing personal health data in a vertically partitioned scenario. *MedInfo*, 264:373–377, 2019.
- [162] Jacqueline Lorene Bender, Alaina B Cyr, Luk Arbuckle, and Lorraine E Ferris. Ethics and privacy implications of using the internet and social media to recruit

- participants for health research: a privacy-by-design framework for online recruitment. *Journal of Medical Internet Research*, 19(4):e7029, 2017.
- [163] Tian Li, Huaqun Wang, Debiao He, and Jia Yu. Blockchain-based privacy-preserving and rewarding private data sharing for iot. *IEEE Internet of Things Journal*, 2022.
- [164] Grant Kelly, Bruce McKenzie, et al. Security, privacy, and confidentiality issues on the internet. *Journal of Medical Internet Research*, 4(2):e861, 2002.
- [165] David Peloquin, Michael DiMaio, Barbara Bierer, and Mark Barnes. Disruptive and avoidable: Gdpr challenges to secondary research uses of data. *European Journal of Human Genetics*, 28(6):697–705, 2020.
- [166] Adrienne C Lahti, Dai Wang, Huiling Pei, Susan Baker, and Vaibhav A Narayan. Clinical utility of wearable sensors and patient-reported surveys in patients with schizophrenia: Noninterventional, observational study. *JMIR mental health*, 8(8):e26234, 2021.
- [167] Lin-Chieh Meng, Shih-Tsung Huang, Ho-Min Chen, Ardeshir Z Hashmi, Fei-Yuan Hsiao, and Liang-Kung Chen. Health care utilization and potentially preventable adverse outcomes of high-need, high-cost middle-aged and older adults: Needs for integrated care models with life-course approach. *Archives of Gerontology and Geriatrics*, 109:104956, 2023.
- [168] Mubeen Akhtar. Innovations in anesthesia delivery: Tailoring care to individual patient needs. *Cosmic Journal of Biology*, 3(1):184–190, 2024.
- [169] Gurleen Kaur and David D Berg. The changing epidemiology of the cardiac intensive care unit. *Critical Care Clinics*, 40(1):1–13, 2024.
- [170] Yong-Feng Ge, Elisa Bertino, Hua Wang, Jinli Cao, and Yanchun Zhang. Distributed cooperative coevolution of data publishing privacy and transparency. *ACM Transactions on Knowledge Discovery from Data*, 18(1):1–23, 2023.
- [171] Jacqueline A De Leeuw, Hetty Woltjer, and Rudolf B Kool. Identification of factors influencing the adoption of health information technology by nurses who are digitally lagging: in-depth interview study. *Journal of medical Internet research*, 22(8):e15630, 2020.
- [172] Jinhyung Lee, Hyeyeong Kim, and Sung J Choi. Do hospital data breaches affect health information technology investment? *DIGITAL HEALTH*, 10:20552076231224164, 2024.

- [173] M Ahmed, E Elaziz, and N Mohamed. Nurse's knowledge, skills, and attitude toward electronic health records. *Journal of Nursing and Health Science*, 9:53–60, 2020.
- [174] Ismail Keshta and Ammar Odeh. Security and privacy of electronic health records: Concerns and challenges. *Egyptian Informatics Journal*, 22(2):177–183, 2021.
- [175] Raag Agrawal and Sudhakaran Prabakaran. Big data in digital healthcare: lessons learnt and recommendations for general practice. *Heredity*, 124(4):525–534, 2020.
- [176] Enrico M Ferrazzi, Luigi Frigerio, Irene Cetin, Patrizia Vergani, Arsenio Spinillo, Federico Prefumo, Edda Pellegrini, and Gianluigi Gargantini. Covid-19 obstetrics task force, lombardy, italy: executive management summary and short report of outcome. *International Journal of Gynecology & Obstetrics*, 149(3):377–378, 2020.
- [177] J Marc Overhage and David McCallie Jr. Physician time spent using the electronic health record during outpatient encounters: a descriptive study. *Annals of internal medicine*, 172(3):169–174, 2020.
- [178] Guy Fagherazzi, Catherine Goetzinger, Mohammed Ally Rashid, Gloria A Aguayo, and Laetitia Huiart. Digital health strategies to fight covid-19 worldwide: challenges, recommendations, and a call for papers. *Journal of Medical Internet Research*, 22(6):e19284, 2020.
- [179] Jalal Halwani and Doris Mouawad. Implementation of e-health innovative technologies in north lebanon hospitals. *Eastern Mediterranean Health Journal*, 27(9): 892–898, 2021.
- [180] Asaph Azaria, Ariel Ekblaw, Thiago Vieira, and Andrew Lippman. Medrec: Using blockchain for medical data access and permission management. In *2016 2nd international conference on open and big data (OBD)*, pages 25–30. IEEE, 2016.
- [181] Kaelan A Moat, Mikayla Wicks, and Michael G Wilson. Citizen brief: Integrating data across sectors for public service improvement in ontario. 2016.
- [182] Daniel J Solove and Woodrow Hartzog. Unifying privacy and data security. 2022.
- [183] Eugenijus Gefenas, J Lekstutiene, V Lukaseviciene, M Hartlev, M Mourby, and K Ó Cathaoir. Controversies between regulations of research ethics and protection of personal data: informed consent at a cross-road. *Medicine, Health Care and Philosophy*, 25(1):23–30, 2022.
- [184] Hameed Hussain Almubarak, Mohamed Khairallah Khouja, and Ahmed Jedidi. Security and privacy recommendation of mobile app for arabic speaking. *International Journal of Electrical & Computer Engineering (2088-8708)*, 12(5), 2022.

- [185] Zhouyu Tian, Lening Qiu, and Litao Wang. Drivers and influencers of blockchain and cloud-based business sustainability accounting in china: Enhancing practices and promoting adoption. *Plos one*, 19(1):e0295802, 2024.
- [186] Alexis Shore, Anisha Reddy, and Carrie Klein. A student-centered privacy model for responsible technology use. *Higher Education Implications for Teaching and Learning During COVID-19*, page 81, 2022.
- [187] Li Zhang, Jianbo Xu, Pandi Vijayakumar, Pradip Kumar Sharma, and Uttam Ghosh. Homomorphic encryption-based privacy-preserving federated learning in iot-enabled healthcare system. *IEEE Transactions on Network Science and Engineering*, 2022.
- [188] Vinden Wylde, Nisha Rawindaran, John Lawrence, Rushil Balasubramanian, Edmond Prakash, Ambikesh Jayal, Imtiaz Khan, Chaminda Hewage, and Jon Platts. Cybersecurity, data privacy and blockchain: a review. *SN Computer Science*, 3(2): 1–12, 2022.
- [189] Phillip Olla, Joseph Tan, Lauren Elliott, and Mustafa Abumeeiz. Security and privacy issues. *Digital Health Care: Perspectives, Applications, and Cases*, page 105, 2022.
- [190] Andrew R Besmer, Jason Watson, and M Shane Banks. Investigating user perceptions of mobile app privacy: An analysis of user-submitted app reviews. *International Journal of Information Security and Privacy (IJISP)*, 14(4):74–91, 2020.
- [191] Lee A Bygrave. The privacy act 1988 (cth): A study in the protection of privacy and the protection of political power. *Federal Law Review*, 19(2):128–153, 1990.
- [192] Jian Guo and Ron Steinfeld. *Advances in Cryptology–ASIACRYPT 2023: 29th International Conference on the Theory and Application of Cryptology and Information Security, Guangzhou, China, December 4–8, 2023, Proceedings, Part I*, volume 14438. Springer Nature, 2024.
- [193] Yifeng Zheng, Menglun Zhou, Songlei Wang, Hejiao Huang, Xiaohua Jia, Xun Yi, and Cong Wang. Secdr: Enabling secure, efficient, and accurate data recovery for mobile crowdsensing. *IEEE Transactions on Dependable and Secure Computing*, 2023.
- [194] Dharmaraj R Patil and Tareek M Pattewar. Majority voting and feature selection based network intrusion detection system. *EAI Endorsed Transactions on Scalable Information Systems*, 9(6):e6–e6, 2022.

- [195] Tehsin Kanwal, Adeel Anjum, and Abid Khan. Privacy preservation in e-health cloud: taxonomy, privacy requirements, feasibility analysis, and opportunities. *Cluster Computing*, 24(1):293–317, 2021.
- [196] Jun Zhao, Kai Zhang, Junqing Gong, and Haifeng Qian. Lavidia: Large-universe, verifiable and dynamic fine-grained access control for e-health cloud. *IEEE Transactions on Information Forensics and Security*, 2024.
- [197] Ashok Kumar Yadav, Karan Singh, Ali H Amin, Laila Almutairi, Theyab R Alsenani, and Ali Ahmadian. A comparative study on consensus mechanism with security threats and future scopes: Blockchain. *Computer Communications*, 201: 102–115, 2023.
- [198] Jingwei Huang, David M Nicol, Rakesh Bobba, and Jun Ho Huh. A framework integrating attribute-based policies into role-based access control. In *Proceedings of the 17th ACM symposium on Access Control Models and Technologies*, pages 187–196, 2012.
- [199] Muhammad Shafay, Raja Wasim Ahmad, Khaled Salah, Ibrar Yaqoob, Raja Jayaraman, and Mohammed Omar. Blockchain for deep learning: review and open challenges. *Cluster Computing*, 26(1):197–221, 2023.
- [200] Rachel Hulkower, Matthew Penn, and Cason Schmit. Privacy and confidentiality of public health information. *Public Health Informatics and Information Systems*, pages 147–166, 2020.
- [201] Isma Masood, Ali Daud, Yongli Wang, Ameen Banjar, and Riad Alharbey. A blockchain-based system for patient data privacy and security. *Multimedia Tools and Applications*, pages 1–25, 2024.
- [202] Rummel John D, Race Margaret S, Horneck, and G the Princeton Workshop Participants. Ethical considerations for planetary protection in space exploration: a workshop, 2012.
- [203] e estonia, 2019. URL <https://e-estonia.com/solutions/healthcare/>. (Accessed 4 July 2019).
- [204] Shubhangi V Urkude, Himanshu Sharma, Seethamsetty Uday Kumar, and Vijaykumar R Urkude. Anatomy of blockchain implementation in healthcare. In *Blockchain Technology: Applications and Challenges*, pages 51–76. Springer, 2021.
- [205] Edgar R Dulce Villarreal, Jose García-Alonso, Enrique Moguel, and Julio Ariel Hurtado Alegría. Blockchain for healthcare management systems: A survey on interoperability and security. *IEEE Access*, 11:5629–5652, 2023.

- [206] Sasidhar Duggineni. Impact of controls on data integrity and information systems. *Science and Technology*, 13(2):29–35, 2023.
- [207] Mesala M Sravani and S Ananiah Durai. Attacks on cryptosystems implemented via vlsi: A review. *Journal of Information Security and Applications*, 60:102861, 2021.
- [208] Mueen Uddin. Blockchain medledger: Hyperledger fabric enabled drug traceability system for counterfeit drugs in pharmaceutical industry. *International Journal of Pharmaceutics*, 597:120235, 2021.
- [209] Mahdi Fahmideh, John Grundy, Aakash Ahmad, Jun Shen, Jun Yan, Davoud Mougouei, Peng Wang, Aditya Ghose, Anuradha Gunawardana, Uwe Aickelin, et al. Engineering blockchain based software systems: Foundations, survey, and future directions. *ACM Computing Surveys*, 2022.
- [210] Johannes Knitza, Rachel Knevel, Karim Raza, Tor Bruce, Ekaterina Eimer, Isabel Gehring, Linda Mathsson-Alm, Maryam Poorafshar, Axel J Hueber, Georg Schett, et al. Toward earlier diagnosis using combined ehealth tools in rheumatology: the joint pain assessment scoring tool (jpast) project. *JMIR mHealth and uHealth*, 8(5):e17507, 2020.
- [211] Guy Aridor, Yeon-Koo Che, and Tobias Salz. The effect of privacy regulation on the data industry: Empirical evidence from gdpr. In *Proceedings of the 22nd ACM Conference on Economics and Computation*, pages 93–94, 2021.
- [212] Naiyana Sahavechaphan, U Suriya, Nattapon Harnsamut, Jessada Phengsuwan, Kamron Aroonrua, et al. An efficient technique for aspect-based ehr access policy administration on abac. In *2011 Ninth International Conference on ICT and Knowledge Engineering*, pages 27–33. IEEE, 2012.
- [213] Ado Adamou Abba Ari, Olga Kengni Ngangmo, Chafiq Titouna, Ousmane Thiare, Alidou Mohamadou, and Abdelhak Mourad Gueroui. Enabling privacy and security in cloud of things: Architecture, applications, security & privacy challenges. *Applied Computing and Informatics*, 20(1/2):119–141, 2024.
- [214] Omar Hasan, Lionel Brunie, and Elisa Bertino. Privacy-preserving reputation systems based on blockchain and other cryptographic building blocks: A survey. *ACM Computing Surveys (CSUR)*, 55(2):1–37, 2022.
- [215] Erikson Júlio De Aguiar, Bruno S Faiçal, Bhaskar Krishnamachari, and Jó Ueyama. A survey of blockchain-based strategies for healthcare. *ACM Computing Surveys (CSUR)*, 53(2):1–27, 2020.

- [216] RC Mesquita and I de Edwards. Systematic literature review of my health record system. *asia pac. J. Health Manag*, 15:14–25, 2020.
- [217] K Selvakumar and S Lokesh. A cryptographic method to have a secure communication of health care digital data into the cloud. *Automatika*, 65(1):373–386, 2024.
- [218] Ariel Ekblaw, Asaph Azaria, John D Halamka, and Andrew Lippman. A case study for blockchain in healthcare: “medrec” prototype for electronic health records and medical research data. In *Proceedings of IEEE open & big data conference*, volume 13, page 13, 2016.
- [219] Ji Yeon Kim. A comparative study of block chain: Bitcoin· namecoin· medibloc. *Journal of Science and Technology Studies*, 18(3):217–255, 2018.
- [220] Bahar Houtan, Abdelhakim Senhaji Hafid, and Dimitrios Makrakis. A survey on blockchain-based self-sovereign patient identity in healthcare. *IEEE Access*, 8: 90478–90494, 2020.
- [221] Bipin Kumar Rai. Pcbhr: patient-controlled blockchain enabled electronic health records for healthcare 4.0. *Health Services and Outcomes Research Methodology*, 23(1):80–102, 2023.
- [222] Guardtime, 2018. URL <https://guardtime.com/blog/world-s-first-blockchain-supported-personal-care-record-platform-launched-by-guardtime-and-partners>. (Accessed 4 July 2019).
- [223] Qi Xia, Emmanuel Boateng Sifah, Kwame Omono Asamoah, Jianbin Gao, Xiaojiang Du, and Mohsen Guizani. Medshare: Trust-less medical data sharing among cloud service providers via blockchain. *IEEE access*, 5:14757–14767, 2017.
- [224] Abdullah Al Omar, Mohammad Shahriar Rahman, Anirban Basu, and Shinsaku Kiyomoto. Medibchain: A blockchain based privacy preserving platform for healthcare data. In *Security, Privacy, and Anonymity in Computation, Communication, and Storage: SpaCCS 2017 International Workshops, Guangzhou, China, December 12-15, 2017, Proceedings 10*, pages 534–543. Springer, 2017.
- [225] Isma Masood, Yongli Wang, Ali Daud, Naif Radi Aljohani, and Hassan Dawood. Towards smart healthcare: patient data privacy and security in sensor-cloud infrastructure. *Wireless Communications and Mobile Computing*, 2018, 2018.
- [226] Alex Roehrs, Cristiano André da Costa, Rodrigo da Rosa Righi, Sandro José Rigo, and Matheus Henrique Wichman. Toward a model for personal health record

- interoperability. *IEEE journal of biomedical and health informatics*, 23(2):867–873, 2018.
- [227] Louisa Walsh, Sophie Hill, Meredith Allan, Susan Balandin, Andrew Georgiou, Isabel Higgins, Ben Kraal, Shaun McCarthy, and Bronwyn Hemsley. A content analysis of the consumer-facing online information about my health record: Implications for increasing knowledge and awareness to facilitate uptake and use. *Health Information Management Journal*, 47(3):106–115, 2018.
- [228] John Bethencourt, Amit Sahai, and Brent Waters. Ciphertext-policy attribute-based encryption. In *2007 IEEE symposium on security and privacy (SP'07)*, pages 321–334. IEEE, 2007.
- [229] Data team M.H. M, 2019. URL <http://www.myhealthmydata.eu/>. (Accessed 1 July 2019).
- [230] MHMD. Shaping our future, newsletter 01, 2018. URL www.myhealthmydata.eu/wp-content/uploads/2017/10/MHMD_newsletter_01_DEF_WEB_pag_doppie_110718.pdf. (Accessed 1 July 2019).
- [231] MHMD. Initial list of main requirements, deliverable 1.1, 2017. URL http://www.myhealthmydata.eu/wp-content/themes/Parallax-One/deliverables/D1.1_InitialList-of-Main-Requirements.pdf.
- [232] Shekha Chenthara, Khandakar Ahmed, Hua Wang, Frank Whittaker, and Zhenxiang Chen. Healthchain: A novel framework on privacy preservation of electronic health records using blockchain technology. *Plos one*, 15(12):e0243043, 2020.
- [233] Orna Fennelly, Dearbhla Moroney, Michelle Doyle, Jessica Eustace-Cook, and Mary Hughes. Key interoperability factors for patient portals and electronic health records: A scoping review. *International Journal of Medical Informatics*, page 105335, 2024.
- [234] Mahyar Amini and Negar Jahanbakhsh Javid. A multi-perspective framework established on diffusion of innovation (doi) theory and technology, organization and environment (toe) framework toward supply chain management system based on cloud computing technology for small and medium enterprises. *Organization and Environment (TOE) Framework Toward Supply Chain Management System Based on Cloud Computing Technology for Small and Medium Enterprises (January 2023)*. *International Journal of Information Technology and Innovation Adoption*, 11:1217–1234, 2023.

- [235] N Venkateswaran and S Prabaharan Prabaharan. An efficient neuro deep learning intrusion detection system for mobile adhoc networks. *EAI Endorsed Transactions on Scalable Information Systems*, 9(6):e7–e7, 2022.
- [236] Swamynathan Ramakrishnan, S Jijitha, and T Amudha. Roadmap of ai and iomt in smart healthcare: Current applications and future perspectives. *Internet of Medical Things in Smart Healthcare*, pages 137–161, 2024.
- [237] Argyro Pountoukidou, Maria Potamiti-Komi, Vrisiis Sarri, Michail Papapanou, Eleni Routsis, Anna Maria Tsiatsiani, Nikolaos Vlahos, and Charalampos Siristatidis. Management and prevention of covid-19 in pregnancy and pandemic obstetric care: a review of current practices. In *Healthcare*, volume 9, page 467. MDPI, 2021.
- [238] A Baker. Crossing the quality chasm: a new health system for the 21st century. *British Medical Journal Publishing Group*, 323(7322):1192–1192, 2001.
- [239] N Jacq. SHARE Roadmap 1: Towards a debate. From Genes to Personalized Healthcare: Grid Solutions for the Life Sciences: Proceedings of HealthGrid, 2007. 126, 164.
- [240] P E Ekmekci. Patients’ rights in cross-border healthcare (Directive 2011/24/EU) and how it applies to Turkey as a negotiating candidate country. *European journal of health law*, 24(4):432–444, 2017.
- [241] A Negrouk, D Horgan, A Gorini, I Cutica, L Leyens, S S Genannt Halfmann, and G Pravettoni. Clinical trials, data protection and patient empowerment in the era of the new EU regulations. *Public health genomics*, 18(6):386–395, 2015.
- [242] B H Rahmouni, T Solomonides, C M Mont, S Shiu, and M Rahmouni. A model-driven privacy compliance decision support for medical data sharing in Europe. *Methods of information in medicine*, 50(04):326–336, 2011.
- [243] H B Rahmouni, T Solomonides, M C Mont, and S Shiu. Privacy compliance and enforcement on European healthgrids: an approach through ontology. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 368:4057–4072, 2010.
- [244] H B Rahmouni, K Munir, M C Mont, and T Solomonides. Semantic generation of clouds privacy policies. In *International Conference on Cloud Computing and Services Science*, pages 15–30. Springer, 2014.
- [245] M Belaazi, H B Rahmouni, and A Bouhoula. Towards a legislation driven framework for access control and privacy protection in public cloud. *11th International Conference on Security and Cryptography (SECRYPT)*, pages 1–6, 2014.

- [246] M Belaazi, H B Rahmouni, and A Bouhoula. An ontology regulating privacy oriented access controls. In *International Conference on Risks and Security of Internet and Systems*, pages 17–35. Springer, 2015.
- [247] I Essefi, H Rahmouni, and M Ladeb. Sensitive Data Discovery in Care Pathways Using Business Process Modelling and HL7-CDA. *International Journal on Advances in Life Sciences*, (1&2):11–11, 2019.
- [248] H B Rahmouni, I Essefi, and M F Ladeb. Enhanced privacy governance in Health Information Systems through business process modelling and HL7. *Procedia Computer Science*, 164:706–713, 2019.
- [249] Y Ding and K Klein. Model-driven application-level encryption for the privacy of e-health data. *2010 International Conference on Availability, Reliability and Security*, pages 341–346, 2010.
- [250] J Li, H Wang, H Jin, and J Yong. Current developments of k-anonymous data releasing. *Proceedings of the National e-Health Privacy and Security Symposium*, pages 109–121, 2006.
- [251] H Wang and Y Zhang. Detection of motor imagery EEG signals employing Naïve Bayes based learning process. *Measurement*, 86:148–158, 2016.
- [252] M Li, S Yu, K Ren, and W Lou. Securing personal health records in cloud computing: Patient-centric and fine-grained data access control in multi-owner settings. In *International conference on security and privacy in communication systems*, pages 89–106. Springer, 2010.
- [253] J Benaloh, M Chase, E Horvitz, and K Lauter. Patient controlled encryption: ensuring privacy of electronic medical records. *Proceedings of the 2009 ACM workshop on Cloud computing security*, pages 103–114, 2009.
- [254] J Huang, M Peng, H Wang, J Cao, W Gao, and X Zhang. A probabilistic method for emerging topic tracking in microblog stream. *World Wide Web*, 20(2):325–350, 2017.
- [255] J Jin, G J Ahn, H Hu, M J Covington, and X Zhang. Patient-centric authorization framework for sharing electronic health records. *Proceedings of the 14th ACM symposium on Access control models and technologies*, pages 125–134, 2009.
- [256] G Ateniese, R Curtmola, B De Medeiros, and D Davis. Medical information privacy assurance: Cryptographic and system aspects. In *International Conference on Security in Communication Networks*, pages 199–218. Springer, 2002.

- [257] M Layouni, K Verslype, M T Sandikkaya, B De Decker, and H Vangheluwe. Privacy-preserving telemonitoring for ehealth. In *IFIP Annual Conference on Data and Applications Security and Privacy*, pages 95–110. Springer, 2009.
- [258] W Tierney. Electronic Health Records: Delivering the Right Information to the Right Health Care Providers at the Right Time, 2011. Available at SSRN 1822397.
- [259] H Wang, J Cao, and Y Zhang. A flexible payment scheme and its role-based access control. *IEEE Transactions on knowledge and Data Engineering*, 17(3):425–436, 2005.
- [260] H Wang, X Jiang, and G Kambourakis. Special issue on Security, Privacy and Trust in network-based Big Data. *Information Sciences-Informatics and Computer Science, Intelligent Systems, Applications: An International Journal*, 318(C):48–50, 2015.
- [261] H Wang, Y Zhang, and J Cao. Effective collaboration with information sharing in virtual universities. *IEEE Transactions on Knowledge and Data Engineering*, 21(6):840–853, 2008.
- [262] X Shi and X Wu. An overview of human genetic privacy. *Annals of the New York Academy of Sciences*, 1387(1):61–61, 2017.
- [263] J L Fernandez-Aleman, I C Senor, P A O Lozoya, and A Toval. Security and privacy in electronic health records: A systematic literature review. *Journal of biomedical informatics*, 46(3):541–562, 2013.
- [264] D Mohammed. US healthcare industry: Cybersecurity regulatory and compliance issues. *Journal of Research in Business, Economics and Management*, 9(5):1771–1776, 2017.
- [265] H Wang and Y Song. Secure cloud-based EHR system using attribute-based cryptosystem and blockchain. *Journal of medical systems*, 42(8):1–9, 2018.
- [266] J Zhang, Y Guo, and Y Chen. Collaborative detection of cybersecurity threats in bigdata. *Int. Arab J. Inf. Technol*, 16(2):186–193, 2019.
- [267] Mikail Mohammed Salim and Jong Hyuk Park. Federated learning-based secure electronic health record sharing scheme in medical informatics. *IEEE Journal of Biomedical and Health Informatics*, 27(2):617–624, 2022.
- [268] Trung Kien Dang, Xiang Lan, Jianshu Weng, and Mengling Feng. Federated learning for electronic health records. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 13(5):1–17, 2022.

- [269] Olivia Choudhury, Aris Gkoulalas-Divanis, Theodoros Salonidis, Issa Sylla, Yoonyoung Park, Grace Hsu, and Amar Das. Differential privacy-enabled federated learning for sensitive health data. *arXiv preprint arXiv:1910.02578*, 2019.
- [270] Eslam Adel, Shaker El-Sappagh, Shaaban Barakat, and Mohammed Elmogy. A unified fuzzy ontology for distributed electronic health record semantic interoperability. In *U-Healthcare Monitoring Systems*, pages 353–395. Elsevier, 2019.
- [271] Joseph Ficek, Wei Wang, Henian Chen, Getachew Dagne, and Ellen Daley. Differential privacy in health research: A scoping review. *Journal of the American Medical Informatics Association*, 28(10):2269–2276, 2021.
- [272] Tengfei Zheng, Yuchuan Luo, Tongqing Zhou, and Zhiping Cai. Towards differential access control and privacy-preserving for secure media data sharing in the cloud. *Computers & Security*, 113:102553, 2022.
- [273] M Chen and M Decary. Artificial Intelligence in healthcare: An essential guide for health leaders. *Healthcare Management Forum*, 33(1):10–18, 2019.
- [274] D Wiljer and Z Hakim. Developing an artificial intelligence-enabled health care practice: Rewiring health care professions for better care. *Journal of Medical Imaging and Radiation Sciences*, 50(4), 2019.
- [275] Z Ahmed, K Mohamed, S Zeeshan, and X Q Dong. Artificial Intelligence with multi-functional machine learning platform development for better healthcare and Precision Medicine. *Database*, 2020, 2020.
- [276] Frank Oemig, Yves Moreau, Ana Ferreira, José Gómez-Pérez, Stefan Decker, Alberto Lluch-Lafuente, and Bruno Lievens. Privacy-preserving machine learning on electronic health records: A systematic review. *Artificial intelligence in medicine*, 102:101757, 2020.
- [277] Gang Luo, Xiaodong Zhang, and Qinghua Li. Exploring machine learning approaches for improving electronic health record confidentiality and security. *Journal of medical systems*, 43(8):250, 2019.
- [278] Michal Cichon and Anwitaman Datta. Machine learning approaches for privacy and security in electronic health records. *IEEE Journal of Biomedical and Health Informatics*, 25(3):812–820, 2021.
- [279] Amalie Dyda, Michael Purcell, Stephanie Curtis, Emma Field, Priyanka Pillai, Kieran Ricardo, Haotian Weng, Jessica C Moore, Michael Hewett, Graham Williams, et al. Differential privacy for public health data: An innovative tool to

- optimize information sharing while protecting data confidentiality. *Patterns*, 2(12), 2021.
- [280] Omar Gutierrez, Jeffreys J Saavedra, Mayra Zurbaran, Augusto Salazar, and Pedro M Wightman. User-centered differential privacy mechanisms for electronic medical records. In *2018 International Carnahan Conference on Security Technology (ICCST)*, pages 1–5. IEEE, 2018.
- [281] Fatemeh Rezaeibagha and Yi Mu. Distributed clinical data sharing via dynamic access-control policy transformation. *International journal of medical informatics*, 89:25–31, 2016.
- [282] ASM Kayes, Wenny Rahayu, Tharam Dillon, and Elizabeth Chang. Accessing data from multiple sources through context-aware access control. 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)*, pages 551–559. IEEE, 2018.
- [283] J. Jin and G. Ahn. Interoperability of electronic health records: A review of technical challenges and standards. *Healthcare Informatics Research*, 24(3):181–187, 2018.
- [284] Huaqun Wang. Anonymous data sharing scheme in public cloud and its application in e-health record. *IEEE Access*, 2018.
- [285] Jinyuan Sun, Xiaoyan Zhu, Chi Zhang, and Yuguang Fang. Hcpp: Cryptography based secure ehr system for patient privacy and emergency healthcare. In *2011 31st International Conference on Distributed Computing Systems*, pages 373–382. IEEE, 2011.
- [286] Hao Jin, Yan Luo, Peilong Li, and Jomol Mathew. A review of secure and privacy-preserving medical data sharing. *IEEE Access*, 7:61656–61669, 2019.
- [287] Mark A Rothstein. The hippocratic bargain and health information technology. *The Journal of Law, Medicine & Ethics*, 38(1):7–13, 2010.
- [288] Devin M Mann, Ji Chen, Rumi Chunara, Paul A Testa, and Oded Nov. Covid-19 transforms health care through telemedicine: evidence from the field. *Journal of the American Medical Informatics Association*, 27(7):1132–1135, 2020.
- [289] Andrew R Watson. Impact of the digital age on transforming healthcare. *Healthcare Information Management Systems: Cases, Strategies, and Solutions*, pages 219–233, 2016.

- [290] Shari M Erickson, Brooke Rockwern, Michelle Koltov, Robert M McLean, Medical Practice, and Quality Committee of the American College of Physicians*. Putting patients first by reducing administrative tasks in health care: a position paper of the american college of physicians. *Annals of internal medicine*, 166(9):659–661, 2017.
- [291] Michael A Tutty, Lindsey E Carlasare, Stacy Lloyd, and Christine A Sinsky. The complex case of ehrs: examining the factors impacting the ehr user experience. *Journal of the American Medical Informatics Association*, 26(7):673–677, 2019.
- [292] Mohamed Abouzahra, Kamran Sartipi, David Armstrong, and Joseph Tan. Integrating data from ehrs to enhance clinical decision making: the inflammatory bowel disease case. In *2014 IEEE 27th International Symposium on Computer-Based Medical Systems*, pages 531–532. IEEE, 2014.
- [293] Ofir Ben-Assuli, Doron Sagi, Moshe Leshno, Avinoah Ironi, and Amitai Ziv. Improving diagnostic accuracy using ehr in emergency departments: A simulation-based study. *Journal of biomedical informatics*, 55:31–40, 2015.
- [294] Bakheet Aldosari. Patients’ safety in the era of emr/ehr automation. *Informatics in Medicine Unlocked*, 9:230–233, 2017.
- [295] Jenny E Han, Marina Rabinovich, Prasad Abraham, Prerna Satyanarayana, T Vivian Liao, Timothy N Udoji, George A Cotsonis, Eric G Honig, and Greg S Martin. Effect of electronic health record implementation in critical care on survival and medication errors. *The American journal of the medical sciences*, 351(6):576–581, 2016.
- [296] Thomas F Osborne, Zachary P Veigulis, David M Arreola, Eliane Rössli, and Catherine M Curtin. Automated ehr score to predict covid-19 outcomes at us department of veterans affairs. *PLoS One*, 15(7):e0236554, 2020.
- [297] Raghavendra Ganiga, Radhika M Pai, and Rajesh Kumar Sinha. Security framework for cloud based electronic health record (ehr) system. *International Journal of Electrical and Computer Engineering*, 10(1):455, 2020.
- [298] Weiran Liu, Xiao Liu, Jianwei Liu, Qianhong Wu, Jun Zhang, and Yan Li. Auditing and revocation enabled role-based access control over outsourced private ehrs. In *2015 IEEE 17th international conference on high performance computing and communications, 2015 IEEE 7th international symposium on cyberspace safety and security, and 2015 IEEE 12th international conference on embedded software and systems*, pages 336–341. IEEE, 2015.

- [299] G Abirami and Revathi Venkataraman. Attribute based access control with trust calculation (abac-t) for decision policies of health care in pervasive environment. *IJITEE*, 8, 2019.
- [300] Evgenia Psarra, Ioannis Patiniotakis, Yiannis Verginadis, Dimitris Apostolou, and Gregoris Mentzas. Securing access to healthcare data with context-aware policies. In *2020 11th International Conference on Information, Intelligence, Systems and Applications (IISA)*, pages 1–6. IEEE, 2020.
- [301] Georgii Kopanitsa. Integration of hospital information and clinical decision support systems to enable the reuse of electronic health record data. *Methods of information in medicine*, 56(4):238–247, 2017.
- [302] Leonidas L Fragidis and Prodromos D Chatzoglou. Implementation of a nationwide electronic health record (ehr): The international experience in 13 countries. *International journal of health care quality assurance*, 31(2):116–130, 2018.
- [303] Timothy McIntosh, Tong Liu, Teo Susnjak, Hooman Alavizadeh, Alex Ng, Raza Nowrozy, and Paul Watters. Harnessing gpt-4 for generation of cybersecurity grc policies: A focus on ransomware attack mitigation. *Computers & Security*, 134: 103424, 2023.
- [304] Ehsan Ghazizadeh, Ebrahim Bagheri, and Pramodini M Singh. Security ontology for electronic health records. *Journal of biomedical informatics*, 53:196–207, 2015.
- [305] Elena Vergara and Javier Lopez. Context-aware attribute-based access control. In *International Conference on Information Security and Cryptology*, pages 165–180. Springer, 2013.
- [306] Jingquan He, Xi Chen, Jianfeng Zhang, and Jingyi Yu. An ontology-driven approach for securing electronic health records. *BMC medical informatics and decision making*, 13(1):12, 2013.
- [307] Han Liu, Shui Yu, and Xiaodong Yang. Ontology-driven context-aware attribute-based access control model for healthcare applications. *Journal of medical systems*, 42(12):249, 2018.
- [308] Derrick Ntalasha, Renfa Li, and Yongheng Wang. Adaptive context-aware design using context state information for the internet of things paradigm. *Journal of Mobile Multimedia*, pages 289–320, 2019.
- [309] Mario Sicuranza and Angelo Esposito. An access control model for easy management of patient privacy in ehr systems. In *8th International Conference for Internet Technology and Secured Transactions (ICITST-2013)*, pages 463–470. IEEE, 2013.

- [310] Marcelo Antonio de Carvalho Junior, Paulo Bandiera-Paiva, et al. Health information system role-based access control current security trends and challenges. *Journal of healthcare engineering*, 2018, 2018.
- [311] Rui Zhang, Ling Liu, and Rui Xue. Role-based and time-bound access and management of ehr data. *Security and communication Networks*, 7(6):994–1015, 2014.
- [312] Angelo Esposito, Mario Sicuranza, and Mario Ciampi. A patient centric approach for modeling access control in ehr systems. In *Algorithms and Architectures for Parallel Processing: 13th International Conference, ICA3PP 2013, Vietri sul Mare, Italy, December 18-20, 2013, Proceedings, Part II 13*, pages 225–232. Springer, 2013.
- [313] Cátia Santos-Pereira, Alexandre B Augusto, Ricardo Cruz-Correia, and Manuel E Correia. A secure rbac mobile agent access control model for healthcare institutions. In *Proceedings of the 26th IEEE international symposium on computer-based medical systems*, pages 349–354. IEEE, 2013.
- [314] Mario Sicuranza, Angelo Esposito, and Mario Ciampi. A view-based access control model for ehr systems. In *Intelligent Distributed Computing VIII*, pages 443–452. Springer, 2015.
- [315] Weiran Liu, Xiao Liu, Jianwei Liu, and Qianhong Wu. Auditing revocable privacy-preserving access control for ehRs in clouds. *The Computer Journal*, 60(12):1871–1888, 2017.
- [316] Karim Abouelmehdi, Abderrahim Beni-Hessane, and Hayat Khaloufi. Big healthcare data: preserving security and privacy. *Journal of big data*, 5(1):1–18, 2018.
- [317] Maryam Zarezadeh, Maede Ashouri Taluki, and Mohammad Siavashi. Attribute-based access control for cloud-based electronic health record (ehr) systems. *ISecure*, 12(2), 2020.
- [318] Aisha Mohammed Alshiky, Seyed M Buhari, and Ahmed Barnawi. Attribute-based access control (abac) for ehr in fog computing environment. *International Journal on Cloud Computing: Services and Architecture (IJCCSA)*, 7(1):27–34, 2017.
- [319] Hao Guo, Wanxin Li, Mark Nejad, and Chien-Chung Shen. Access control for electronic health records with hybrid blockchain-edge architecture. In *2019 IEEE International Conference on Blockchain (Blockchain)*, pages 44–51. IEEE, 2019.
- [320] Redwan Walid, Karuna P Joshi, and Seung Geol Choi. Semantically rich differential access to secure cloud ehr. In *2023 IEEE 9th 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), pages 1–9. IEEE, 2023.
- [321] Kwangsoo Seol, Young-Gab Kim, Euijong Lee, Young-Duk Seo, and Doo-Kwon Baik. Privacy-preserving attribute-based access control model for xml-based electronic health record system. *IEEE Access*, 6:9114–9128, 2018.
- [322] Litun Patra, Udai Pratap Rao, Pooja Choksi, and Akhil Chaurasia. Controlling access to ehealth data using request denial cache in xacml reference architecture for abac. In *2022 IEEE 3rd Global Conference for Advancement in Technology (GCAT)*, pages 1–8. IEEE, 2022.
- [323] Amel Arfaoui, Soumaya Cherkaoui, Ali Kribeche, Sidi Mohammed Senouci, and Mohamed Hamdi. Context-aware adaptive authentication and authorization in internet of things. In *ICC 2019-2019 IEEE International Conference on Communications (ICC)*, pages 1–6. IEEE, 2019.
- [324] Rayane El Sibai, Nader Gemayel, Jacques Bou Abdo, and Jacques Demerjian. A survey on access control mechanisms for cloud computing. *Transactions on Emerging Telecommunications Technologies*, 31(2):e3720, 2020.
- [325] Lingfeng Chen and Doan B Hoang. Novel data protection model in healthcare cloud. In *2011 IEEE International Conference on High Performance Computing and Communications*, pages 550–555. IEEE, 2011.
- [326] S Padmapriya, R Shankar, R Thiagarajan, S Arun, BS Liya, and B Gunasundari. Preserving privacy scheme using data-caac mechanism in e-health based on hybrid edge computing. In *2021 3rd International Conference on Advances in Computing, Communication Control and Networking (ICAC3N)*, pages 1394–1399. IEEE, 2021.
- [327] ASM Kayes, Jun Han, and Alan Colman. Ontcaac: an ontology-based approach to context-aware access control for software services. *The Computer Journal*, 58(11):3000–3034, 2015.
- [328] Mohammad H Yarmand, Kamran Sartipi, and Douglas G Down. Behavior-based access control for distributed healthcare environment. In *2008 21st IEEE International Symposium on Computer-Based Medical Systems*, pages 126–131. IEEE, 2008.
- [329] Mohammad H Yarmand, Kamran Sartipi, and Douglas G Down. Behavior-based access control for distributed healthcare systems. *Journal of Computer Security*, 21(1):1–39, 2013.

- [330] Changbo Ke, Jiayu Wu, Fu Xiao, Zhiqiu Huang, and Yunfei Meng. A privacy risk assessment scheme for fog nodes in access control system. *IEEE Transactions on Reliability*, 71(4):1513–1526, 2021.
- [331] Mario Sicuranza and Mario Ciampi. A semantic access control for easy management of the privacy for ehr systems. In *2014 Ninth International Conference on P2P, Parallel, Grid, Cloud and Internet Computing*, pages 400–405. IEEE, 2014.
- [332] Jorge Calvillo-Arbizu, Isabel Román-Martínez, and Laura M Roa-Romero. Standardized access control mechanisms for protecting iso 13606-based electronic health record systems. In *IEEE-EMBS International Conference on Biomedical and Health Informatics (BHI)*, pages 539–542. IEEE, 2014.
- [333] Sharad Dixit, Karuna P Joshi, and Seung Geol Choi. Multi authority access control in a cloud ehr system with ma-abe. In *2019 IEEE international conference on edge computing (EDGE)*, pages 107–109. IEEE, 2019.
- [334] Redwan Walid, Karuna P Joshi, Seung Geol Choi, and Dae-young Kim. Cloud-based encrypted ehr system with semantically rich access control and searchable encryption. In *2020 IEEE International Conference on Big Data (Big Data)*, pages 4075–4082. IEEE, 2020.
- [335] Mor Peleg, Dizza Beimel, Dov Dori, and Yaron Denekamp. Situation-based access control: Privacy management via modeling of patient data access scenarios. *Journal of Biomedical Informatics*, 41(6):1028–1040, 2008.
- [336] Dizza Beimel, Mor Peleg, and Tim Redmond. Reasoning about access-control situations with owl. In *The 11th Intl Protégé Conference, Amsterdam, Netherlands*, 2009.
- [337] Xiao Dong, Reza Samavi, and Thodoros Topaloglou. Coc: An ontology for capturing semantics of circle of care. *Procedia Computer Science*, 63:589–594, 2015.
- [338] Enamul Kabir and Hua Wang. Conditional purpose based access control model for privacy protection. volume 92, pages 137–144, 01 2009.
- [339] Xiaoxun Sun, Hua Wang, Jiuyong Li, and Yanchun Zhang. Injecting purpose and trust into data anonymisation. *Computers & Security*, 30:332–345, 07 2011. doi: 10.1016/j.cose.2011.05.005.
- [340] Ji Zhang, Xiaohui Tao, and Hua Wang. Outlier detection from large distributed databases. *World Wide Web*, 17, 07 2014. doi: 10.1007/s11280-013-0218-4.

- [341] Meriem Zerkouk, Paulo Cavalcante, Abdallah Mhamed, Jerome Boudy, and Belhadri Messabih. Behavior and capability based access control model for personalized telehealthcare assistance. *Mobile Networks and Applications*, 19:392–403, 2014.
- [342] Enamul Kabir, Abdun Mahmood, Hua Wang, and Abdul Mustafa. Microaggregation sorting framework for k-anonymity statistical disclosure control in cloud computing. *IEEE Transactions on Cloud Computing*, PP:408–417, 08 2015. doi: 10.1109/TCC.2015.2469649.
- [343] Enamul Kabir. A role-involved purpose-based access control model. *Information Systems Frontiers*, 14:809–822, 07 2012.
- [344] Hua Wang, Yanchun Zhang, Jinli Cao, and Vijay Varadharajan. Achieving secure and flexible m-services through tickets. *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on*, 33:697 – 708, 12 2003. doi: 10.1109/TSMCA.2003.819917.
- [345] Jiao Yin, Mingjian Tang, Jinli Cao, and Hua Wang. Apply transfer learning to cybersecurity: Predicting exploitability of vulnerabilities by description. *Knowledge-Based Systems*, 210, 10 2020. doi: 10.1016/j.knosys.2020.106529.
- [346] Abiodun Muyideen Mustapha, Temitope Elizabeth Abioye, Olusanya Oyedele, Folasade Mercy Okikiola, and Christianah Yetunde Alonge. A systematic literature review of ontology-based techniques in medical diagnosis. *Available at SSRN 4394368*.
- [347] Kamal Sharma, Sandeep Gupta, Rajbir Kaur, and Manoj Kumar. Ontology driven electronic health record. In *2016 International Conference on Computing, Communication and Automation (ICCCA)*, pages 940–944. IEEE, 2016.
- [348] Anne M Tall and Cliff C Zou. A framework for attribute-based access control in processing big data with multiple sensitivities. *Applied Sciences*, 13(2):1183, 2023.
- [349] Parveen Dhillon and Manpreet Singh. An extended ontology model for trust evaluation using advanced hybrid ontology. *Journal of Information Science*, page 01655515221128424, 2023.
- [350] Lena Wahlberg. Legal ontology, scientific expertise and the factual world. *Journal of Social Ontology*, 3(1):49–65, 2017.
- [351] Yip Chi Kiong, Sellappan Palaniappan, and Nor Adnan Yahaya. Health ontology system. In *2011 7th International Conference on Information Technology in Asia*, pages 1–4. IEEE, 2011.

- [352] Eric Helms and Laurie Williams. Evaluating access control of open source electronic health record systems. In *Proceedings of the 3rd workshop on software engineering in health care*, pages 63–70, 2011.
- [353] Yifan Yang, Run-hua Shi, Kunchang Li, Zhiwei Wu, and Shuhao Wang. Multiple access control scheme for ehrs combining edge computing with smart contracts. *Future Generation Computer Systems*, 129:453–463, 2022.
- [354] Bharath Chintagunta, Namit Katariya, Xavier Amatriain, and Anitha Kannan. Medically aware gpt-3 as a data generator for medical dialogue summarization. In *Machine Learning for Healthcare Conference*, pages 354–372. PMLR, 2021.
- [355] Diane M Korngiebel and Sean D Mooney. Considering the possibilities and pitfalls of generative pre-trained transformer 3 (gpt-3) in healthcare delivery. *NPJ Digital Medicine*, 4(1):93, 2021.
- [356] Timothy R McIntosh, Teo Susnjak, Tong Liu, Paul Watters, and Malka N Halgamuge. From google gemini to openai q*(q-star): A survey of reshaping the generative artificial intelligence (ai) research landscape. *arXiv preprint arXiv:2312.10868*, 2023.
- [357] Yassi Moghaddam, Heather Yurko, Haluk Demirkan, Nathan Tymann, and Ammar Rayes. *The future of work: how artificial intelligence can augment human capabilities*. business expert press, 2020.
- [358] Karen Yeung. A study of the implications of advanced digital technologies (including ai systems) for the concept of responsibility within a human rights framework. *MSI-AUT (2018)*, 5, 2018.
- [359] Heike Felzmann, Eduard Fosch-Villaronga, Christoph Lutz, and Aurelia Tamò-Larrieux. Towards transparency by design for artificial intelligence. *Science and Engineering Ethics*, 26(6):3333–3361, 2020.
- [360] Cristina Brandão, Guilhermina Rego, Ivone Duarte, and Rui Nunes. Social responsibility: a new paradigm of hospital governance? *Health Care Analysis*, 21: 390–402, 2013.
- [361] Alejo Jose G Sison, Marco Tulio Daza, Roberto Gozalo-Brizuela, and Eduardo C Garrido-Merchán. Chatgpt: More than a weapon of mass deception, ethical challenges and responses from the human-centered artificial intelligence (hcai) perspective. *arXiv preprint arXiv:2304.11215*, 2023.
- [362] Liam G McCoy, Connor TA Brenna, Stacy S Chen, Karina Vold, and Sunit Das. Believing in black boxes: machine learning for healthcare does not need explainability to be evidence-based. *Journal of clinical epidemiology*, 142:252–257, 2022.

- [363] Timothy R McIntosh, Tong Liu, Teo Susnjak, Paul Watters, Alex Ng, and Malka N Halgamuge. A culturally sensitive test to evaluate nuanced gpt hallucination. *IEEE Transactions on Artificial Intelligence*, 1(01):1–13, 2023.
- [364] Jessica Davis. 2020 saw a sharp increase in ransomware attacks on health-care, 2020. URL <https://healthitsecurity.com/news/2020-saw-a-sharp-increase-in-ransomware-attacks-on-healthcare>.
- [365] Priti Tagde, Sandeep Tagde, Tanima Bhattacharya, Pooja Tagde, Hitesh Chopra, Rokeya Akter, Deepak Kaushik, and Md Habibur Rahman. Blockchain and artificial intelligence technology in e-health. *Environmental Science and Pollution Research*, 28:52810–52831, 2021.
- [366] H Wang and Y Song. Secure cloud-based ehr system using attribute-based cryptosystem and blockchain. *Journal of medical systems*, 42(8):1–9, 2018.
- [367] MC Beuscart-Zéphir, E Borycki, P Carayon, MWM Jaspers, and S Pelayo. Evolution of human factors research and studies of health information technologies: the role of patient safety. *Yearbook of medical informatics*, 22(01):67–77, 2013.
- [368] Bart Custers, Alan M Sears, Francien Dechesne, Iilina Georgieva, Tommaso Tani, and Simone Van der Hof. *EU personal data protection in policy and practice*. Springer, 2019.
- [369] Y Wang, A Zhang, P Zhang, and H Wang. Cloud-assisted ehr sharing with security and privacy preservation via consortium blockchain. *IEEE Access*, 7:136704–136719, 2019.
- [370] G G Dagher, J Mohler, M Milojkovic, and P B Marella. Ancile: Privacy-preserving framework for access control and interoperability of electronic health records using blockchain technology. *Sustainable cities and society*, 39:283–297, 2018.
- [371] E Zaghloul, T Li, and J Ren. Security and privacy of electronic health records: decentralized and hierarchical data sharing using smart contracts. *2019 International Conference on Computing, Networking and Communications (ICNC)*, pages 375–379, 2019.
- [372] A Kayes, W Rahayu, T Dillon, E Chang, and J Han. Context-aware access control with imprecise context characterization for cloud-based data resources. *Future Generation Computer Systems*, 93:237–255, 2019.
- [373] D Akarca, P Xiu, D Ebbitt, B Mustafa, H Al-Ramadhani, and A Albey-Atti. Blockchain secured electronic health records: Patient rights, privacy and cybersecurity. *2019 10th International Conference on Dependable Systems, Services and Technologies (DESSERT)*, pages 108–111, 2019.

- [374] A H A Omar. The Effect of Electronic Health Records On Undergraduate and Postgraduate Medical Education: A Scoping Review, 2019.
- [375] N Kshetri. Blockchain's roles in strengthening cybersecurity and protecting privacy. *Telecommunications policy*, 41(10):1027–1038, 2017.
- [376] F Rezaeibagha and Y Mu. Distributed clinical data sharing via dynamic access-control policy transformation. *International journal of medical informatics*, 89:25–31, 2016.
- [377] Lina Zhu, Zuochang Zhang, Guoen Xia, and Caoqing Jiang. Research on vulnerability ontology model. In *2019 IEEE 8th Joint International Information Technology and Artificial Intelligence Conference (ITAIC)*, pages 657–661. IEEE, 2019.
- [378] C S Kruse, B Smith, H Vanderlinden, and A Nealand. Security techniques for the electronic health records. *Journal of medical systems*, 41(8):1–9, 2017.
- [379] S Tanwar, K Parekh, and R Evans. Blockchain-based electronic health-care record system for healthcare 4.0 applications. *Journal of Information Security and Applications*, 50:102407–102407, 2020.
- [380] M Farhadi, H Haddad, and H Shahriar. Static analysis of hipaa security requirements in electronic health record applications. *2018 IEEE 42nd Annual Computer Software and Applications Conference (COMPSAC)*, 2:474–479, 2018.
- [381] P Vimalachandran, Y Zhang, J Cao, L Sun, and J Yong. Preserving data privacy and security in australian my health record system: A quality health care implication. *International Conference on Web Information Systems Engineering*, pages 111–120, 2018.
- [382] A Spencer and S Patel. Applying the data protection act 2018 and general data protection regulation principles in healthcare settings. *Nursing Management*, 26(1), 2019.
- [383] S F Malamed. *Sedation-e-book: a guide to patient management*. Elsevier Health Sciences, 2017.
- [384] W Varndell, A Hodge, and M Fry. Triage in australian emergency departments: Results of a new south wales survey. *Australasian Emergency Care*, 22(2):81–86, 2019.
- [385] A S M Kayes, J Han, and A Colman. An ontology-based approach to context-aware access control for software services in WISE, pp. 410–420, 2013.

- [386] A Kayes, J Han, W Rahayu, T Dillon, M S Islam, and A Colman. A policy model and framework for context-aware access control to information resources. *The Computer Journal*, 62(5):670–705, 2019.
- [387] Ann Cavoukian. Privacy by design: the definitive workshop. a foreword by ann cavoukian, ph. d. *Identity in the Information Society*, 3(2):247–251, 2010.
- [388] Emmanuel Eze, Rob Gleasure, and Ciara Heavin. Mobile health solutions in developing countries: a stakeholder perspective. *Health Systems*, 9(3):179–201, 2020.
- [389] Linda W Peute, Gaby-Anne Wildenbos, Thomas Engelsma, Blake J Lesselroth, Valentina Lichtner, Helen Monkman, David Neal, Lex Van Velsen, Monique W Jaspers, and Romaric Marcilly. Overcoming challenges to inclusive user-based testing of health information technology with vulnerable older adults: Recommendations from a human factors engineering expert inquiry. *Yearbook of Medical Informatics*, 31(01):074–081, 2022.
- [390] Wencheng Sun, Zhiping Cai, Yangyang Li, Fang Liu, Shengqun Fang, and Guoyan Wang. Security and privacy in the medical internet of things: a review. *Security and Communication Networks*, 2018:1–9, 2018.
- [391] Sharon Silow-Carroll, Jennifer N Edwards, and Diana Rodin. Using electronic health records to improve quality and efficiency: the experiences of leading hospitals. *Issue Brief (Commonw Fund)*, 17(1):40, 2012.
- [392] Ming Li, Shucheng Yu, Kui Ren, and Wenjing Lou. Securing personal health records in cloud computing: Patient-centric and fine-grained data access control in multi-owner settings. In *Security and Privacy in Communication Networks: 6th International ICST Conference, SecureComm 2010, Singapore, September 7-9, 2010. Proceedings 6*, pages 89–106. Springer, 2010.
- [393] A Kayes, W Rahayu, P Watters, M Alazab, T Dillon, and E Chang. Achieving security scalability and flexibility using fog-based context-aware access control. *Future Generation Computer Systems*, 107.
- [394] Google Colab. Welcome to colaboratory, 2021. <https://research.google.com/colaboratory,2021>. Accessed: 2021-10.
- [395] Q Mamun. A conceptual framework of personally controlled electronic health record (pcehr) system to enhance security and privacy. *International Conference on Applications and Techniques in Cyber Security and Intelligence*, pages 304–314, 2017.

- [396] S Samet, M T Ishraque, and A Sharma. Privacy-preserving personal health record (p3hr) a secure android application. *Proceedings of the 7th International Conference on Software and Information Engineering*.
- [397] Kelly Caine and Rima Hanania. Patients want granular privacy control over health information in electronic medical records. *Journal of the American Medical Informatics Association*, 20(1):7–15, 2013.
- [398] A Baker. Crossing the quality chasm: a new health system for the 21st century. *British Medical Journal Publishing Group*, 323, 2001.
- [399] N Jacq et al. Share roadmap 1: Towards a debate, 2007. From Genes to Personalized Healthcare: Grid Solutions for the Life Sciences: Proceedings of HealthGrid 2007, vol. 126, p. 164.
- [400] T Grandison, C Johnson, and J Kiernan. Hippocratic databases: current capabilities and future trends,” in Handbook of database security, 2008. pp. 409–429, Springer.
- [401] C M Johnson and T Grandison. Compliance with data protection laws using hippocratic database active enforcement and auditing. *IBM systems journal*, 46(2):255–264, 2007.
- [402] A Zhang and X Lin. Towards secure and privacy-preserving data sharing in e-health systems via consortium blockchain. *Journal of medical systems*, 42(8):1–18, 2018.
- [403] T. T Kuo, H. E Kim, and L Ohno-Machado. Blockchain distributed ledger technologies for biomedical and health care applications. *Journal of the American Medical Informatics Association*, 24(6):1211–1220, 2017.
- [404] C Esposito, A Santis, G Tortora, H Chang, and K.-K R Choo. Blockchain: A panacea for healthcare cloud-based data security and privacy? *IEEE Cloud Computing*, 5(1):31–37, 2018.
- [405] Salvatore F Pileggi. Ontology in Hybrid Intelligence: a concise literature review, 2023.
- [406] Kais Tahar, Tamara Martin, Raphael Yonglimou, Holm Verbuecheln, Dagmar Graessner, and Krefting. Rare Diseases in Hospital Information Systems-An Interoperable Methodology for Distributed Data Quality Assessments. *Methods of Information in Medicine AAM*, 2023.

- [407] José Blanco, Bruno Miguel, and Tomášpitner Rossi. A Comparative Study of Energy Domain Ontologies. In *Web Information Systems and Technologies: 16th International Conference*, volume 2020, pages 43–58. Springer International Publishing, 2020.
- [408] Joseph Awotunde, Bamidele, Olaiyafolorunsho, Olayinka Isah Olawale Mustapha, Mulikat Olufunmilayo Olusanya, Kazeem Bola Akanbi, and Moses Abiodun. An Enhanced Internet of Things Enabled Type-2 Fuzzy Logic for Healthcare System Applications. In *Recent Trends on Type-2 Fuzzy Logic Systems: Theory, Methodology and Applications*, pages 133–151. Springer International Publishing, 2023.
- [409] Wei Chen, Theresia Avila Tong, and Bria. A Review of Ontology-Based Safety Management in Construction. *Sustainability*, 15(1):413–413, 2023.
- [410] Parveen Dhillon and Manpreet Singh. An extended ontology model for trust evaluation using advanced hybrid ontology. *Journal of Information Science*, pages 01655515221128424–01655515221128424, 2023.
- [411] Kulsoom S Bughio and Leslie F Sikos. Knowledge Organization Systems to Support Cyber-Resilience in Medical Smart Home Environments. In *Cybersecurity for Smart Cities*, pages 2023–2023. Springer.
- [412] Carine Souveyet and Rébeccadeneckère. Design and Modeling in Pervasive Information Systems. In *The Evolution of Pervasive Information Systems*, pages 19–41. Springer International Publishing, 2023.
- [413] Melissa Zimdars, Megan E Cullinan, and Kilhoe Na. Alternative health groups on social media, misinformation, and the (de) stabilization of ontological security. *New Media & Society*, pages 14614448221146171–14614448221146171, 2023.
- [414] Suniti Purbey, Brijesh Khandelwal, and Ashutosh Kumar Choudhary. Design of a blockchain based secure and efficient ontology generation model for multiple data genres using augmented stratification in healthcare industry, 2023.
- [415] Sánchez-Zas, Víctor A Carmen, Mario Villagrà, Xavier Vega-Barbas, José Ignacio Larriva-Novo, Julio Moreno, and Berrocal. Ontology-based approach to real-time risk management and cyber-situational awareness. *Future Generation Computer Systems*, 141:462–472, 2023.
- [416] Michael Debellis and Biswanath Dutta. From ontology to knowledge graph with agile methods: the case of COVID-19 CODO knowledge graph. *International Journal of Web Information Systems ahead-of-print*, 2022.

- [417] Zahra Rezaei and Mohammad H Vahidnia. Effective medical center finding during COVID-19 pandemic using a spatial DSS centered on ontology engineering. *GeoJournal*, pages 1–15, 2022.
- [418] Flora Amato, Valentina Casola, Giovanni Cozzolino, Alessandra De Benedictis, Nicola Mazzocca, and Francesco Moscato. A security and privacy validation methodology for e-health systems. *ACM Transactions on Multimedia Computing*, 17(2s):1–22, 2021.
- [419] Saeedeh Ghanadbashi, Zahra Safavifar, Farshad Taebi, and Fatemeh Golpayegani. Handling uncertainty in self-adaptive systems: an ontology-based reinforcement learning model. *Journal of Reliable Intelligent Environments*, pages 1–26, 2023.
- [420] Azra Bashir, Renuka Nagpal, Deepti Mehrotra, and Manju Bala. Interoperability of Electronic Health Records for Dyslipidemia using Knowledge graphs. *2023 13th International Conference on Cloud Computing, Data Science & Engineering (Confluence)*, pages 124–129, 2023.
- [421] Min Yang, Xingshu Chen, Liuyan Tan, Xiao Lan, and Yonggang Luo. Listen carefully to experts when you classify data: A generic data classification ontology encoded from regulations. *Information Processing & Management*, 60(2):103186–103186, 2023.
- [422] Anil Sharma and Suresh Kumar. Ontology-based semantic retrieval of documents using Word2vec model. *Data & Knowledge Engineering*, 144:102110–102110, 2023.
- [423] Dylan Mcgagh, Anant Jani, John Williams, Harshana Liyanage, Uy Hoang, Cecilia Okusi, Julian Sherlock, Filipa Ferreira, Simon Ivelinayonova, and De Lusignan. A novel ontological approach to track social determinants of health in primary care. In *Measuring Ontologies for Value Enhancement: Aligning Computing Productivity with Human Creativity for Societal Adaptation: First International Workshop, MOVE 2020, Virtual Event*, pages 227–240. Springer, 2020.
- [424] Frank Wawrzik, Khushnood Adil Rafique, Farin Rahman, and Christoph Grimm. *Ontology Learning Applications of Knowledge Base Construction for Microelectronic Systems Information*, 14:176–176, 2023.
- [425] Miroslav Zarić, Sašarsovski, Petarvasiljević Brankomarkoski, and Veliborpremčevski. An Approach to the Semantic Representation of the Local Government Strategic Planning Process: Ontology-Driven Simulation Method for Assessing Economic Impacts. *Applied Sciences*, 13(3):1258–1258, 2023.
- [426] Gerald Ovono and Sihlemoyo. Goal Modeling for Linked Data Exploitation of Municipalities Data Access in South Africa. In *Software Engineering Application*

- in Systems Design: Proceedings of 6th Computational Methods in Systems and Software 2022*, volume 1, pages 234–248. Springer International Publishing, 2023.
- [427] Yuhang Zhou, Tengfei Bao, Xiaosong Shu, Yueyang Li, and Yangtao Li. BIM and ontology-based knowledge management for dam safety monitoring. *Automation in Construction*, 145:104649–104649, 2023.
- [428] Francisco M Garcia-Moreno, Maria Bermudez-Edo, José Manuel Pérez, José Luis Mármol, and María José Rodríguez-Fórtiz Garrido. A Conceptual Model of Health Monitoring Systems Centered on ADLs Performance in Older Adults. In *Advances in Conceptual Modeling: ER 2022 Workshops, CMLS, EmpER, and JUSMOD*, pages 25–34. Springer International Publishing, 2022.
- [429] Dewei Yang, Shuai Liu, Jinbao Sheng, Xuehui Peng, and Huiwen Wang. Construction and implementation of ontology model of dam break emergency plan. *Advances in Measurement Technology, Disaster Prevention and Mitigation*, pages 213–220, 2023.
- [430] Hamed Barangi, Shekoufeh Rahimi, Bahman Zamani, and Hossein Moradi. An Ontology-based Approach to Facilitate Semantic Interoperability of Context-Aware Systems. *International Computer Conference*, 2023.
- [431] Nguyen Thuan, Duy Hoang, Hoanh-Su Dang-Pham, Prasanta Le, Tuan Q Bhattacharya, and Phan. Introduction to Information Systems Research in Vietnam: A Shared Vision. In *Information Systems Research in Vietnam*, pages 2023–2023. Springer.
- [432] Khairul Prawira, Djarothindarto Thamrin, and Ekoindrajit. *Application of Enterprise Architecture in Digital Transformation of Insurance Companies*, pages 856–865, 2023.
- [433] Jing Qian and Yi Liu. Quantitative scenario construction of typical disasters driven by ontology data. *Journal of Safety Science and Resilience*, 4(2):159–166, 2023.
- [434] Rosario Minardi, Maria Luisa Villani, and Antonio De Nicola. Semantic Reasoning for Geolocalized Assessment of Crime Risk in Smart Cities. *Smart Cities*, 6(1):179–195, 2023.
- [435] Gerald Ovono and Sihlemoyo. Conceptual Linked Data Model for South African Municipalities Public Services Domain. In *Data Science and Algorithms in Systems: Proceedings of 6th Computational Methods in Systems and Software 2022*, volume 2, pages 197–208. Springer International Publishing, 2023.

- [436] Philipp Kernstock, Leonard Przybilla, Jason Thatcher, and Helmut Krcmar. Can't Get No Satisfaction?"-The Case for Broadening Information Systems Research on E-Commerce, 2023.
- [437] Jose Alvarez-Rodríguez, Roy María, Eduardo Mendieta, Juan Cibrián, and Llorens. Towards a method to quantitatively measure toolchain interoperability in the engineering lifecycle: a case study of digital hardware design. *Computer Standards & Interfaces*, pages 103744–103744, 2023.
- [438] Patrick Lambrix. Database and Web Information Systems Group: Publications, 2023.
- [439] Alexandra I Khalyasmaa, Alina I Stepanova, A Stanislav, Pavel V Eroshenko, and Matrenin. Review of the Digital Twin Technology Applications for Electrical Equipment Lifecycle Management. *Mathematics*, 11(6):1315–1315, 2023.
- [440] De Matos, Eduardo Everton, Ramãotiburski Viegas, and Fabiano Hessel. Context-Aware Security in the Internet of Things: A Review. In *Advanced Information Networking and Applications: Proceedings of the 37th International Conference on Advanced Information Networking and Applications (AINA-2023)*, volume 3, pages 518–531. Springer International Publishing, 2023.
- [441] Hnin Nu Lwin, Prattanapunnakitikashem Nu, and Trinthananusak. E-Health Research in Southeast Asia: A Bibliometric Review. *Sustainability*, 15(3):2559–2559, 2023.
- [442] Reval Puneeth, Govindaswamy Prabhu, and Parthasarathy. A Survey on Security and Interoperability of Electronic Health Records Sharing Using Blockchain Technology. *Acta Informatica Pragensia*, 12(1), 2023.
- [443] Santhosh Kumar, K S, J Hanumanthappa, S P Shiva Prakash, and Kirill Krinkin. Relationship-Based AES Security Model for Social Internet of Things. In *Intelligent Systems and Applications: Select Proceedings of ICISA 2022*, pages 143–151. Springer Nature, 2023.
- [444] Mehdi Gheisari, Hamid Esmacilinajafabadi, Jafar A Alzubi, Jiechao Gao, Guojun Wang, Aaqifazaal Abbasi, and Aniello Castiglione. OBPP: An ontology-based framework for privacy-preserving in IoT-based smart city. *Future Generation Computer Systems*, 123:1–13, 2021.
- [445] Touré, Philip Vasundra, Kristin Krauss, Gnodtke, Deepak Jaschabuchhorn, Unni, Jean Louis Petarhorki, and Raisaro. FAIRification of health-related data using semantic web technologies in the Swiss Personalized Health Network. *Scientific Data*, 10(1):127–127, 2023.

- [446] Ozonze, Philip J Obinwa, Adrian A Scott, and Hopgood. Automating Electronic Health Record Data Quality Assessment. *Journal of Medical Systems*, 47(1):23–23, 2023.
- [447] Ozonze, Philip J Obinwa, Adrian A Scott, and Hopgood. Automating Electronic Health Record Data Quality Assessment. *Journal of Medical Systems*, 47(1):23–23, 2023.
- [448] Prosper Yeng, Adam Kandabongee, Bian Szekeres, Einar Arthur Yang, and Snekkenes. Mapping the psychosocialcultural aspects of healthcare professionals’ information security practices: Systematic mapping study. *JMIR human factors*, 8(2):17604–17604, 2021.
- [449] Gabriele Gäbler, Deborah Lycett, and Walter Gall. Integrating a New Dietetic Care Process in a Health Information System: A System and Process Analysis and Assessment. *International Journal of Environmental Research and Public Health*, 19(5):2491–2491, 2022.
- [450] Ines Meriah and Latifa Ben Arfarabai. Comparative study of ontologies based iso 27000 series security standards. *Procedia Computer Science*, 160:85–92, 2019.
- [451] Lamia Moudoubah, Khalifa Mansouri, and Mohammed Qbadou. COBIT 5 Concepts: Towards the Development of an Ontology Model. *Advances in Information, Communication and Cybersecurity: Proceedings of ICI2C’21*, pages 247–256, 2022.
- [452] Stephen V Flowerday and Christos Xenakis. Security and Privacy in Distributed Health Care Environments. *Methods of information in medicine*, 61:1–002, 2022.
- [453] Ines Meriah and Latifa Ben Arfarabai. Comparative study of ontologies based iso 27000 series security standards. *Procedia Computer Science*, 160:85–92, 2019.
- [454] Van Rossem, Annalisa Wouter, and Pelizza. The ontology explorer: A method to make visible data infrastructures for population management. *Big Data & Society*, 9(1):20539517221104087–20539517221104087, 2022.
- [455] Kamrun Nahar and Asif Qumer Gill. Integrated identity and access management metamodel and pattern system for secure enterprise architecture. *Data & Knowledge Engineering*, 140:102038–102038, 2022.
- [456] Ebtsam Adel, Shaker El-Sappagh, Sherif Barakat, and Mohammed Elmogy. Ontology-based electronic health record semantic interoperability: A survey. *U-healthcare monitoring systems*, pages 315–352, 2019.
- [457] Salvatore F Pileggi. Ontology in Hybrid Intelligence: a concise literature review, 2023.

- [458] Maryam Ahmadi, Abbas Sheikhtaheri, Maryam Foziyeh Tahmasbi, Fatemeh Es-lamijahromi, and Rangrazjeddi. A competency framework for Ph. D. programs in health information management. *International Journal of Medical Informatics*, 168:104906–104906, 2022.
- [459] Farnaz Mohammadi, Angelikipanou, Eirini Christoforosntantogian, Emmanouil-panaousis Karapistoli, and Christos Xenakis. CUREX: seCUre and pRivatehEalth data eXchange. *IEEE/WIC/ACM International Conference on Web Intelligence-Companion Volume*, pages 263–268, 2019.
- [460] Nafei Zhu, Baocun Chen, Siyu Wang, Da Teng, and Jingsha He. Ontology-based approach for the measurement of privacy disclosure. *Information Systems Frontiers*, 24(5):1689–1707, 2022.
- [461] Farnaz Mohammadi, Angelikipanou, Eirini Christoforosntantogian, Emmanouil-panaousis Karapistoli, and Christos Xenakis. CUREX: seCUre and pRivatehEalth data eXchange. *IEEE/WIC/ACM International Conference on Web Intelligence-Companion Volume*, pages 263–268, 2019.
- [462] Jorsiele Cerqueira. An Ontology for Context-aware Middleware for Dependable Medical Systems. *Proceedings of the 11th Latin-American Symposium on Dependable Computing*, pages 79–83, 2022.
- [463] Ronald Ojino, Luisa Mich, and Nereymvungi. Hotel room personalization via ontology and rule-based reasoning. *International Journal of Web Information Systems ahead-of-print*, 2022.
- [464] Tian Bai, Lan Huang, Shuyu Guo, Yichen Liu, Minfei Wu, Guishan Gu, and Xiao Luo. Integrating knowledge from Case Report: a medical-ontology based multimodal information system with structured summary, 2022.
- [465] Luković, Vanja, Danijela Sašačuković, Goran Milošević, and Devedžić. An ontology-based module of the information system ScolioMedIS for 3D digital diagnosis of adolescent scoliosis. *Computer Methods and Programs in Biomedicine*, 178:247–263, 2019.
- [466] Zijie Ren, Jianhua Shi, and Muhammad Imran. Data evolution governance for ontology-based digital twin product lifecycle management. *IEEE Transactions on Industrial Informatics*, 19(2):1791–1802, 2022.
- [467] Carlo Sansone and Giancarlo Sperlí. Legal Information Retrieval systems: State-of-the-art and open issues. *Information Systems*, 106:101967–101967, 2022.

- [468] Tiago F Pereira, Arthur Matta, Carlos M Mayea, Frederico Pereira, Nelson Monroy, João Jorge, and Tiago Rosa. A web-based Voice Interaction framework proposal for enhancing Information Systems user experience. *Procedia Computer Science*, 196:235–244, 2022.
- [469] Zahra Rezaei and Mohammad H Vahidnia. Effective medical center finding during COVID-19 pandemic using a spatial DSS centered on ontology engineering. *GeoJournal*, pages 1–15, 2022.
- [470] Zahra Rezaei and Mohammad H Vahidnia. Effective medical center finding during COVID-19 pandemic using a spatial DSS centered on ontology engineering. *GeoJournal*, pages 1–15, 2022.
- [471] Eman M Abounassar, Passent El-Kafrawy, and Ahmed A El-Latif. Security and interoperability issues with internet of things (IoT) in healthcare industry: A survey. *Security and Privacy Preserving for IoT and 5G Networks: Techniques, Challenges, and New Directions*, pages 159–189, 2022.
- [472] Oluwasefunmi T Arogundade, Adebayo Abayomi-Alli, and Sanjay Misra. An ontology-based security risk management model for information systems. *Arabian Journal for Science and Engineering*, 45:6183–6198, 2020.
- [473] Elena Doynikova, Andrey Fedorchenko, and Igor Kotenko. A semantic model for security evaluation of information systems. *Journal of Cyber Security and Mobility*, pages 301–330, 2020.
- [474] Fouzia F Ozair, Nayer Jamshed, Amit Sharma, and Praveen Aggarwal. Ethical issues in electronic health records: A general overview. *Perspectives in clinical research*, 6(2):73, 2015.
- [475] M Kumaar, D Samiayya, P M Vincent, K Srinivasan, C. Y. Chang, and H Ganesh. A hybrid framework for intrusion detection in healthcare systems using Deep Learning. *Frontiers in Public Health*, 9:2022–2022.
- [476] Jobie Budd, Benjamin S Miller, Erin M Manning, Vasileios Lampos, Mengdie Zhuang, Michael Edelstein, Geraint Rees, Vincent C Emery, Molly M Stevens, Neil Keegan, et al. Digital technologies in the public-health response to covid-19. *Nature medicine*, 26(8):1183–1192, 2020.
- [477] Stephen J Mooney and Vikas Pejaver. Big data in public health: terminology, machine learning, and privacy. *Annual review of public health*, 39:95–112, 2018.
- [478] Z Ahmed. Practicing precision medicine with intelligently integrative clinical and multi-omics Data Analysis. *Human Genomics*, 14(1), 2020.

- [479] Vangalur Alagar, Alaa Alsaig, Olga Ormandjiva, and Kaiyu Wan. Context-based security and privacy for healthcare iot. In *2018 IEEE International Conference on Smart Internet of Things (SmartIoT)*, pages 122–128. IEEE, 2018.
- [480] Paul R Demuro and Carolyn Petersen. Managing privacy and data sharing through the use of health care information fiduciaries. In *Context Sensitive Health Informatics: Sustainability in Dynamic Ecosystems*, pages 157–162. IOS Press, 2019.
- [481] Victoria Kisekka, Justin Scott Giboney, et al. The effectiveness of health care information technologies: evaluation of trust, security beliefs, and privacy as determinants of health care outcomes. *Journal of medical Internet research*, 20(4): e9014, 2018.
- [482] Margaret F A Otlowski. Disclosing genetic information to at-risk relatives: new australian privacy principles, but uniformity still elusive. *The Medical Journal of Australia*, 202(6):335–337, 2015.
- [483] M Chen and M Decary. Artificial Intelligence in healthcare: An essential guide for health leaders. *Healthcare Management Forum*, 33(1):10–18, 2019.
- [484] W Abramson, A J Hall, P Papadopoulos, N Pitropakis, and W J Buchanan. A distributed trust framework for privacy-preserving machine learning, 2020. Trust, Privacy and Security in Digital Business, pp. 205–220.
- [485] Tanzir Ul Islam, Reza Ghasemi, and Noman Mohammed. Privacy-preserving federated learning model for healthcare data. In *2022 IEEE 12th Annual Computing and Communication Workshop and Conference (CCWC)*, pages 0281–0287. IEEE, 2022.
- [486] Waldemar W Koczkodaj, Mirosław Mazurek, Dominik Strzałka, Alicja Wolny-Dominiak, and Marc Woodbury-Smith. Electronic health record breaches as social indicators. *Social Indicators Research*, 141:861–871, 2019.
- [487] J Wang, M Li, Y He, H Li, K Xiao, and C Wang. A blockchain based privacy-preserving incentive mechanism in crowdsensing applications. *Ieee Access*, 6:17545–17556, 2018.
- [488] K Fan, S Wang, Y Ren, H Li, and Y Yang. Medblock: Efficient and secure medical data sharing via blockchain. *Journal of medical systems*, 42(8):1–11, 2018.
- [489] K N Griggs, O Ossipova, C P Kohlios, A N Baccarini, E A Howson, and T Haya-jneh. Healthcare blockchain system using smart contracts for secure automated remote patient monitoring. *Journal of medical systems*, 42(7):1–7, 2018.

- [490] A Kayes, W Rahayu, and T Dillon. An ontology-based approach to dynamic contextual role for pervasive access control. *2018 IEEE 32nd International Conference on Advanced Information Networking and Applications (AINA)*, pages 601–608, 2018.
- [491] Jestine Paul, Meenatchi Sundaram Muthu Selva Annamalai, William Ming, Ahmad Al Badawi, Bharadwaj Veeravalli, and Khin Mi Mi Aung. Privacy-preserving collective learning with homomorphic encryption. *IEEE Access*, 9:132084–132096, 2021.
- [492] Aderonke Justina Ikuomola and Oluremi O Arowolo. Securing patient privacy in e-health cloud using homomorphic encryption and access control. *International Journal of Computer Networks and Communications Security*, 2(1):15–21, 2014.
- [493] AM Vengadapurvaja, G Nisha, R Aarth, and N Sasikaladevi. An efficient homomorphic medical image encryption algorithm for cloud storage security. *Procedia computer science*, 115:643–650, 2017.
- [494] Jafar A Alzubi, Omar A Alzubi, Majdi Beseiso, Anil Kumar Budati, and K Shankar. Optimal multiple key-based homomorphic encryption with deep neural networks to secure medical data transmission and diagnosis. *Expert Systems*, 39(4):e12879, 2022.
- [495] V Subramaniaswamy, V Jagadeeswari, V Indragandhi, Rutvij H Jhaveri, V Vijayakumar, Ketan Kotecha, and Logesh Ravi. Somewhat homomorphic encryption: Ring learning with error algorithm for faster encryption of iot sensor signal-based edge devices. *Security and Communication Networks*, 2022, 2022.
- [496] Desam Vamsi and Pradeep Reddy. Electronic health record security in cloud: medical data protection using homomorphic encryption schemes. In *Research Anthology on Securing Medical Systems and Records*, pages 853–877. IGI Global, 2022.
- [497] Jon L Mills and Kelsey Harclerode. Privacy, mass intrusion, and the modern data breach. *Fla. L. Rev.*, 69:771, 2017.
- [498] Amel Alghrani, Margaret Brazier, Anne-Maree Farrell, Danielle Griffiths, and Neil Allen. Healthcare scandals in the nhs: crime and punishment. *Journal of Medical Ethics*, 37(4):230–232, 2011.
- [499] Peter Schaar. Privacy by design. *Identity in the Information Society*, 3(2):267–274, 2010.
- [500] R Sreejith and S Senthil. Dynamic Data Infrastructure Security for interoperable e-healthcare systems: A semantic feature-driven NoSQL intrusion attack detection model. *BioMed Research International*, 2022:1–26, 2022.

- [501] B D Deebak, F H Memon, X Cheng, K Dev, J Hu, S A Khowaja, N M Qureshi, and K H Choi. Seamless privacy-preservation and Authentication Framework for IOT-enabled Smart eHealth Systems. *Sustainable Cities and Society*, 80:103661–103661, 2022.
- [502] A Sharma and S Kumar. Machine learning and ontology-based novel semantic document indexing for information retrieval. *Computers & Industrial Engineering*, 176:108940–108940, 2023.
- [503] J A Fries, E Steinberg, S Khattar, S L Fleming, J Posada, A Callahan, and N H Shah. Ontology-driven weak supervision for clinical entity classification in Electronic Health Records. *Nature Communications*, 12(1), 2021.
- [504] S S Sahoo, K Kobow, J Zhang, J Buchhalter, M Dayyani, D P Upadhyaya, K Prantzalos, M Bhattacharjee, I Blumcke, S Wiebe, and S D Lhatoo. Ontology-based feature engineering in machine learning workflows for Heterogeneous Epilepsy Patient Records. *Scientific Reports*, 12(1):2022–2022.
- [505] N Zhu, B Chen, S Wang, D Teng, and J He. Ontology-based approach for the measurement of privacy disclosure. *Information Systems Frontiers*, 24(5):1689–1707, 2021.
- [506] E Yehia, H Boshnak, S Abdelgaber, A Abdo, and D S Elzanfaly. Ontology-based clinical information extraction from physician’s free-text notes. *Journal of Biomedical Informatics*, 98:103276–103276, 2019.
- [507] A D Bosco, R Vieira, B Zanotto, and A P Da Silva Etges. Ontology based classification of electronic health records to support value-based health care, 2021. *Intelligent Systems*, pp. 359-371.
- [508] Deborah L McGuinness, Frank Van Harmelen, et al. Owl web ontology language overview. *W3C recommendation*, 10(10):2004, 2004.
- [509] Alex Wang and Kyunghyun Cho. Bert has a mouth, and it must speak: Bert as a markov random field language model. *arXiv preprint arXiv:1902.04094*, 2019.
- [510] Ekaba Bisong and Ekaba Bisong. Google colab. *Building machine learning and deep learning models on google cloud platform: a comprehensive guide for beginners*, pages 59–64, 2019.
- [511] Eugene J Schweitzer. Reconciliation of the cloud computing model with us federal electronic health record regulations. *Journal of the American Medical Informatics Association*, 19(2):161–165, 2012.

- [512] Clemens Scott Kruse, Benjamin Frederick, Taylor Jacobson, and D. Kyle Monticone. Cybersecurity in healthcare: A systematic review of modern threats and trends. *Technology and Health Care*, 25(1):1–10, 2017.
- [513] Jin Li, Amir Talaei-Khoei, Holly Seale, Pradeep Ray, and C. Raina Macintyre. Health care provider adoption of ehealth: Systematic literature review. *Interactive Journal of Medical Research*, 4(2):e7, 2015.
- [514] Yun Zhang and Min Yang. Intelligent cloud storage usage for electronic health record system. *Journal of Medical Systems*, 41(3):44, 2017.
- [515] Vassiliki Papakonstantinou, Mikaela Poulymenopoulou, Flora Malamateniou, and George Vassilacopoulos. Access control for cloud-based emergency medical data management systems. *Health Informatics Journal*, 22(4):812–824, 2016.
- [516] Shahrouz Ghanbari and Mohammad Abdollahi Azgomi. A taxonomy and survey of cloud resource orchestration techniques. *ACM Computing Surveys (CSUR)*, 51(3):1–34, 2018.
- [517] Nasser Bagheri, Harjinder Singh, Chris Nugent, Pradeep Gaur, Vahid H Tafreshi, Liliana Jaqueline Villalba, and Juan Carlos Augusto. A systematic literature review on security and privacy of electronic health record systems: technical perspectives. *Health informatics journal*, 22(2):68–78, 2016.
- [518] MyungYeon Kim, Kevin B Johnson, John Derk, Dan Groves, Norman Kwon, Peter Szolovits, and Boyang Xie. Electronic health records: A guide to ehr selection, implementation and incentives. *Journal of industrial engineering and management*, 9(2):309–329, 2016.
- [519] Amanda Johnson, Michael Smith, and Jane Doe. Exploring ontology-based security in electronic health record systems. *Journal of Healthcare Informatics Research*, 4(2):123–145, 2020.
- [520] Patricia Groves, Reem Kayyali, David Knott, and Stephanie Van Kuiken. Digital health ethics and the responsible innovation of artificial intelligence. *Digital health*, 6:2055207620939045, 2020.
- [521] Jaehyuk Han, Younghoon Kim, Jonghwa Park, and Minjoo Kang. Evaluating the security and privacy risks of blockchain-based electronic health record systems. *Journal of Biomedical Informatics*, 125:104801, 2022.
- [522] Emily Parker, Thomas Brown, and Laura Green. Evaluating privacy ontologies for electronic health records. *Health Informatics Journal*, 27(1):25–39, 2021.

- [523] Robert Anderson, Elizabeth Taylor, and Kevin Liu. Ethical considerations in ehr security and privacy. *Healthcare Ethics*, 6(3):201–215, 2020.
- [524] Mustafa Nazmi Aydin and Ammar Ali. Ethical and security challenges in electronic health records: A review. *Journal of medical systems*, 45(8):90, 2021.
- [525] William R Shadish, Thomas D Cook, and Donald T Campbell. Experimental and quasi-experimental designs for generalized causal inference. 2021.
- [526] Yutong Yao, Huixin Wang, and Yuanlong Li. A novel security and privacy-preserving scheme for electronic health record systems. *IEEE Access*, 9:70524–70536, 2021.
- [527] Sanjay Gupta, Asad Basheeruddin, and Pardeep Kumar. A systematic review on information security risks and threats in healthcare information systems. *Computers in Biology and Medicine*, 118:103641, 2020.
- [528] Haruko Ando, Masaki Ohkubo, and Kazunori Ikeda. Information security governance and management in healthcare: A systematic literature review. *International Journal of Medical Informatics*, 157:104608, 2022.
- [529] William M.K. Trochim and James P. Donnelly. *The Research Methods Knowledge Base*. Cengage, 4th edition, 2020.
- [530] David Blumenthal. Health information technology—what is its role in patient safety? *New England Journal of Medicine*, 382(13):1185–1187, 2020.
- [531] Ebtsam Adel, Shaker El-Sappagh, Sherif Barakat, Jong-Wan Hu, and Mohammed Elmogy. An extended semantic interoperability model for distributed electronic health record based on fuzzy ontology semantics. *Electronics*, 10(14):1733, 2021.
- [532] Zubair Afzal, Martijn J Schuemie, Jan C van Blijderveen, Elif F Sen, Miriam CJM Sturkenboom, and Jan A Kors. Improving sensitivity of machine learning methods for automated case identification from free-text electronic medical records. *BMC medical informatics and decision making*, 13(1):1–11, 2013.
- [533] Tao Gu, Xiao Hang Wang, Hung Keng Pung, and Da Qing Zhang. An ontology-based context model in intelligent environments. *arXiv preprint arXiv:2003.05055*, 2020.
- [534] Colin A Puri, Karthik Gomadam, Prateek Jain, Peter Z Yeh, and Kunal Verma. Multiple ontologies in healthcare information technology: Motivations and recommendation for ontology mapping and alignment. In *ICBO*, 2011.

- [535] Abdul Quamar, Fatma Özcan, Dorian Miller, Robert J Moore, Rebecca Niehus, and Jeffrey Kreulen. Conversational bi: an ontology-driven conversation system for business intelligence applications. *Proceedings of the VLDB Endowment*, 13(12):3369–3381, 2020.
- [536] Alexey Tsymbal, Sonja Zillner, and Martin Huber. Ontology-supported machine learning and decision support in biomedicine. In *Data Integration in the Life Sciences: 4th International Workshop, DILS 2007, Philadelphia, PA, USA, June 27-29, 2007. Proceedings 4*, pages 156–171. Springer, 2007.
- [537] Lei Zhang, Faming Qi, Zeyu Wang, Enhong Wang, and Zhen Liu. Integrating semantic knowledge to tackle zero-shot text classification. *arXiv preprint arXiv:1911.04841*, 2019.
- [538] Pascal Hitzler, Adila A. Krisnadhi, and Krzysztof Janowicz. *Ontology engineering with ontology design patterns: Foundations and applications*. IOS Press, 2016.
- [539] Surya Sharma, Ali Reza Alebouyeh, Sanath Perera, May D. Barreto, and Samarath Kumar Udgata. The role of ontologies for sustainable, semantically interoperable and trustworthy ehr solutions. *Healthcare technology letters*, 7(1):14–22, 2020.
- [540] A. Stevens and A. Dehghan. Using machine learning to identify disease-relevant genes. *Current opinion in genetics & development*, 50:48–53, 2018.
- [541] David Sánchez and Montserrat Batet. C-sanitized: A privacy model for document redaction and sanitization. *Journal of the Association for Information Science and Technology*, 67(1):148–163, 2016.
- [542] Hisham Kanaan, Khalid Mahmood, and Varun Sathyan. An ontological model for privacy in emerging decentralized healthcare systems. In *2017 IEEE 13th International Symposium on Autonomous Decentralized System (ISADS)*, pages 107–113. IEEE, 2017.