



A Cryptographic and Threat Modeling Approach for Privacy-Compliant Digital Forensics in Healthcare Systems.

Submitted by ,

Jyothi Avulareddygar

Supervisor : Dipo Dunsin

Student ID : 2436065

Project Submitted as part of the requirements for the award of MSc in Cyber
Security and Digital Forensic May 2026

Date of Submission: 08/06/2026

Declaration

I, Jyothi Avulareddygar, declare that I am the sole author of this Project; that all references cited have been consulted; that I have conducted all work of which this is a record, and that the finished work lies within the prescribed word limits. This work has not previously been accepted as part of any other degree submission.

Signed : Jyothi Avulareddygar

Date : 08/06/2026

Form Of Consent

I, **Jyothi Avulareddygar** hereby consent that my Project, submitted in candidature for the MSc Cybersecurity and Digital Forensics, if successful, may be made available for inter-library loan or photocopying (subject to the law of copyright), and that the title and abstract may be made available to outside organisations.

Signed : Jyothi Avulareddygar

Date : 08/06/2026

Abstract

This dissertation aims to address the issue of patient privacy in healthcare settings and its clash with digital forensics. As healthcare organisations rely more and more on digital forensics to handle cyber incidents involving electronic health records, audit logs and connected medical devices, traditional forensic methods like full-disk imaging and unscoped log collection often include data for patients that weren't part of the incident. This results in additional privacy violations that are consistently in violation of the HIPAA minimum necessary requirement and the GDPR data minimisation and purpose limitation principles, and introduces a governance gap that can't be addressed with existing forensic readiness and maturity models.

To overcome this, the study introduces a Privacy-Preserving Healthcare Forensics (PPHF) framework which incorporates cryptographic privacy protection directly into the forensic investigation lifecycle in healthcare settings. The four-phase framework is designed to address NIST Privacy Framework functions Identify-P, Control-P, and Protect-P, as well as to counteract the privacy threats outlined in LINDDUN: Linkability, Identifiability, Disclosure, and Non-compliance. A Design Science Research methodology is followed while the design and evaluation of a proof-of-concept prototype on synthetic EHR and audit log datasets is carried out, which showcases field-level AES based encryption, SSE style search over encrypted audit logs, CP-ABE style policy evaluation and hash chained audit logging.

Evaluation with LINDDUN shows that PPHF can substantially mitigate the secondary privacy risks, as it encrypts the PHI fields, and enables analysis only on encrypted indexes, and only decrypts fields relevant to the analysis of the incident based on role, case and legal attributes, with a sufficient forensic capability for scoped investigations. The alignment of the PPHF components with NIST Privacy Framework illustrates the design's adherence to established privacy risk management practices and how it will enable organisations to achieve HIPAA/GDPR compliance by providing technical evidence of minimum-necessary access and accountability. The dissertation concludes that cryptographically integrated, threat-modelled forensic lifecycles like PPHF are a viable direction towards privacy-compliant digital forensics

in healthcare and presents recommendations and future work that will allow for transition from proof-of-concept to real-world deployment.

Keywords: privacy-preserving digital forensics; healthcare; electronic health records; field-level encryption; searchable symmetric encryption; attribute-based encryption; LINDDUN; NIST Privacy Framework; HIPAA; GDPR; Design Science Research.

Table of Content

1 Introduction.....	8
1.1 Problem Statement.....	10
1.2 Aim & Objectives.....	10
1.3 Dissertation Structure.....	11
2 Literature Review.....	14
2.1 Digital Forensics in Healthcare Environments.....	14
2.2 Regulatory Frameworks Governing Healthcare Forensics.....	16
2.3 Cryptographic Privacy-Enhancing Technologies.....	17
2.4 Privacy Threat Modelling in Healthcare.....	19
2.5 Access Control and Audit in Healthcare Forensics.....	20
2.6 Synthesis and Identified Research Gap.....	21
3 Methodology.....	23
3.1 Research Philosophy: Positivism.....	23
3.2 Research Approach: Deductive Reasoning.....	24
3.3 Research Strategy: Design Science Research.....	24
3.4 The DSRM Process Model and Dissertation Mapping.....	25
3.5 Methodological Coherence.....	26
3.6 Data Strategy: Synthetic Datasets.....	27
3.7 Ethical Considerations.....	27
3.8 Methodological Limitations.....	28
4 Framework Design.....	29
4.1 Introduction to the PPHF Framework.....	29
4.2 Multi-Phase Forensic Lifecycle Architecture.....	29
4.2.1 Phase 1 — Privacy-Aware Evidence Identification.....	29
4.2.2 Phase 2 — Encrypted Evidence Acquisition.....	30
4.2.3 Phase 3 — Privacy-Preserving Analysis.....	31
4.3 Access Control, Selective Partial Decryption, and Phase 4.....	31
4.3.1 CP-ABE Access Control Model.....	31
Figure 3: DFD representation of CP -ABE Model.....	32
4.3.2 Selective Partial Decryption and Reporting.....	32
4.4 Audit and Chain-of-Custody Layer.....	33
4.4.1 NIST Privacy Framework and Auditability.....	33
4.5 Regulatory Alignment Summary.....	34
5 Implementation.....	36
5.1 Implementation Environment and Technology Stack.....	36
5.2 Synthetic Dataset Construction.....	37

5.3 Field-Level Encryption Implementation.....	38
5.4 Hashing for Chain-of-Custody.....	39
5.5 Searchable Symmetric Encryption over Encrypted Logs.....	40
5.6 Selective Partial Decryption Logic.....	41
5.7 End-to-End Forensic Scenario.....	42
5.8 Implementation Challenges and Trade-offs.....	43
6 Evaluation.....	44
6.1 Evaluation Design and Criteria.....	44
6.2 LINDDUN-Based Threat Evaluation.....	45
6.2.1 Linkability and Identifiability.....	45
6.2.2 Non-repudiation and Detectability.....	46
6.2.3 Disclosure, Unawareness, and Non-compliance.....	47
6.3 NIST Privacy Framework Alignment.....	47
6.4 Forensic Capability vs. Privacy Trade-off.....	49
6.5 Methodological Reflection: DSR Evaluation.....	50
7 Conclusions and Future Work.....	52
7.1 Synthesis of Objectives and Research Findings.....	52
7.2 Evidence-Based Recommendations.....	54
7.3 Limitations of the Research.....	55
7.4 Future Work.....	56
8 Personal Reflection.....	58
9 References.....	61

List of Tables & Figures

List of Tables

Table	Title	Page
Table 1	DSRM Phases vs Activity vs Dissertation	21
Table 2	Regulatory alignment summary	39
Table 3	summary classification	43
Table 4	PPHF component vs NIST function	45

List of Figures

Figure	Title	Page
Figure 1	DFD representation of Phase 1	27
Figure 2	DFD presentation of phrase 2	28
Figure 3	DFD representation of CP-ABE Model	29
Figure 4	DFD representation of NIST Privacy framework	30
Figure 5	code snippet for environment implementation	33
Figure 6	code snippet for the synthetic dataset construction	34
Figure 7	code snippet for audit log construction	35
Figure 8	code snippet for field level encryption	38
Figure 9	code snippet for synthetic HR data set encryption	38
Figure 10	hashing chain of custody	36
Figure 11	code snippet for searchable symmetric encryption	37
Figure 12	code snippet for searchable symmetric encryption	38
Figure 13	Partial decryption logic	38
Figure 14	investigator attribution	39

Figure 15	end to end forensic scenario	40
-----------	------------------------------	----

1 Introduction

Healthcare is a key target for cybercrime and attackers are increasingly targeting electronic health records (EHRs), clinical applications that are cloud-hosted, and networked medical devices to exfiltrate or extort sensitive patient data [10], [14], [94]. Digital evidence from these systems is now crucial to the understanding of how incidents happen, to mitigate threats and to provide due diligence to regulators and patients [21] [80]. At the same time, health care information is especially private, including details about health, treatment and life circumstances, and is therefore regulated with special laws, such as the Health Insurance Portability and Accountability Act (HIPAA) in the United States and the General Data Protection Regulation (GDPR) in the European Union [8] and [18]. These regimes require that any processing of protected health information (PHI) – including for security investigations – be based on the principles of data minimisation, purpose limitation and minimum necessary use and disclosure of patient records [22] [55].

Most of the conventional digital forensic methodologies were developed for general IT and criminal investigation scenarios, with the focus being on retaining as much potentially relevant information as possible in a forensically sound way [26], [74]. Methods like broad network traffic capture, unscoped log collection, and full-disk imaging are appreciated because they provide the greatest opportunity for later analysis of the artefacts that might be of interest [43, 91]. In a healthcare context, however, they frequently end up ingesting records of patients not necessarily associated with the incident being investigated, which are termed secondary privacy violations: harms that are not necessarily at the site of the breach or attack, but are committed during the forensic response [31] [40] [52]. Practical guidance for HIPAA-compliant cloud forensics recognises that most tools do not natively support the selective acquisition of data, or privacy-aware analysis at the level of individual records or fields, despite the clear warnings in HIPAA's minimum necessary rule and GDPR's data minimisation principle [12], [82]. This typically puts healthcare organisations into a false dichotomy between thorough investigation and privacy compliance and privacy-aware investigation and lack of knowledge about the incidents [30], [59].

In the interim, privacy and risk management frameworks have developed and provide a structured approach to thinking about and resolving such tensions at design time [9, 34]. The NIST Privacy Framework Version 1.0 is a risk- and outcome-based framework that includes functions like Identify-P, Control-P and Protect-P, which can help organisations identify their data flows, have robust access controls and incorporate privacy into enterprise risk decisions [34], [37]. A few privacy threat modelling techniques, such as LINDDUN, classify privacy threats into Linkability, Identifiability, Disclosure of information, and Non-compliance risks and offer systematic “threat trees” to identify and mitigate privacy risks in the system data flows [23] [35]. However, these instruments are seldom used in the digital forensic lifecycle, but on the operational systems and services, which leaves the investigation processes that are consuming the data of these operational systems as unmodelled, privacy-blind spaces [64]. This separation can result in organisations with well-defined privacy programs to inadvertently disrupt their own controls during incident response, because forensic operations are not necessarily as well-designed as production operations [50, 92].

In the meantime, cryptographic-based privacy enhancing technologies have been developed to afford much more granular control over what an adversary can see, and when and under what conditions, even in the presence of adversaries [29, 66]. Field-level encryption can be used to keep certain information fields encrypted at rest and in transit, but still be able to process the information in a meaningful way (such as names or diagnoses) [49], [86]. Searchable symmetric encryption (SSE) allows to search for keywords or patterns over encrypted logs and records without decrypting all the data, thus decreasing the unnecessary exposure of data during analysis [76] [84]. Selective partial decryption, especially with attribute-based encryption (ABE), allows decryption to be tied to an attribute like the role of an investigator, the case identifier and legal authorisation, which means it is possible to implement data minimisation and purpose limitation at the cryptographic level in addition to organisational policy [41], [72], [77]. While these advances have been achieved, none of the current frameworks specifically consider these technologies and systematically integrate them into a multi-phase forensic architecture for healthcare specifically designed to address LINDDUN and aligned with the NIST Privacy Framework [46], [90].

1.1 Problem Statement

Patient data is a uniquely sensitive asset, and it is one of the most targeted assets by cybercriminals, especially in the healthcare sector[5]. If a security incident does happen, investigators are caught in a delicate balancing act: they want to find the digital evidence but they have to do so within the legal and ethical boundaries of safeguarding patients' privacy whose records aren't relevant to the security incident being investigated[8]. In reality, in the process of conducting a forensic investigation, this tension can lead to an inadvertent secondary privacy violation, and thus to a clash with existing regulatory frameworks like HIPAA and GDPR[12]. To tackle this gap, a paradigm change in the approach to digital forensics is needed that focuses on recovering meaningful evidence of the incident while also protecting the privacy of non-participating patients.

1.2 Aim & Objectives

To develop and evaluate a complete Privacy-Preserving Healthcare Forensics (PPHF) framework that integrates cryptographic privacy preserving principles into the digital forensic investigation process in healthcare environment and to validate the fundamental ideas of the developed framework by implementing a proof-of-concept prototype which enables privacy compliant evidence acquisition and analysis without revealing unconcerned patient information.

Objectives

1. To examine the current state of the art in digital forensics and privacy-enhancing technologies, healthcare data protection laws, and the key challenges that the current state of the art in digital forensics does not address, and to provide a theoretical and regulatory basis to support the proposed PPHF framework.
2. To articulate the research philosophy, design approach and architectural principles that will guide the proposed PPHF framework, provide a multi-phase structure, access control model and data minimisation mechanisms via design artefacts, and thus create a blueprint for privacy-compliant forensic investigation in clinical settings.

3. To build a prototype of the core PPHF framework principles — field-level encryption of healthcare records, privacy-preserving search over encrypted audit logs, and selective partial decryption limited to evidence relevant to the incident — that can be used in the field to validate the proposed design.
4. To compare the proposed framework with known security and privacy threat models and test its resistance to known attack vectors and adherence to HIPAA, GDPR and NIST Privacy Framework standards to assess its viability as a privacy-compliant forensic solution for the healthcare environment.
5. To reflect on the trade-offs between forensic investigative capability and patient privacy protection, and to identify opportunities for future research and real-world deployment of the findings from the framework proposal, prototype implementation, and evaluation to come to evidence-informed conclusions and actionable recommendations.

1.3 Dissertation Structure

Chapter 1: Introduction

This chapter provides background to the research by examining the tension between the requirements of digital forensic investigations and the responsibilities for patient privacy in healthcare settings. It is a formal statement of the problem, statement of the aim and five objectives, and a general outline of the dissertation.

Chapter 2: Literature Review

In this chapter, we review the literature on digital forensic methodologies, privacy-enhancing cryptographic technologies, and data protection laws in healthcare such as HIPAA and GDPR, and critically assess them. It pinpoints the gaps in the traditional forensic practice that do not adequately respect patient confidentiality, and provides a theoretical basis for the proposed PPHF framework.

Chapter 3: Research Methodology

This chapter provides the justification of the research philosophy, approach and research strategy that was adopted in this study, which was based on a positivist, deductive research philosophy and design science research strategy. It provides reasons for the suitability of this methodology for the proposal and validation of a technical framework in a regulated healthcare environment.

Chapter 4: Design Process

The chapter describes the architecture of the proposed PPHF framework in a multi-phase manner, along with the access control model, data minimisation mechanisms, and interactions between the cryptographic components, using suitable design artefacts like UML diagrams and data flow diagrams. It creates an entire blueprint of the structure prior to any implementation.

Chapter 5: Implementation

This chapter explains the development of the proof-of-concept prototype to show the feasibility of the core PPHF framework principles on synthetic healthcare datasets. It outlines important implementation choices, technical issues faced, and explains the deliberate scope limitation between the proposed complete framework and the working prototype.

Chapter 6: Evaluation of Results

This chapter reviews and analyzes the proposed PPHF framework and prototype based on structured threat models (LINDDUN) and compliance with NIST Privacy Framework. It translates the findings and identifies the effectiveness of the framework in ensuring patient privacy while maintaining the ability to conduct forensic investigations

Chapter 7: Conclusions and Recommendations

This chapter brings together the results of all the previous chapters, showing how each objective was achieved, and provides recommendations, based on the evidence found, for the practical implementation of privacy-preserving forensics in healthcare. It also pinpoints the gaps in the current work and suggests future research and framework extension.

Chapter 8: Reflection

This chapter shares a personal reflection of the research process, highlighting lessons learned in the areas of time management, academic research skills and technical development during the dissertation process. It is about personal development and learning experience, not about the technical aspects of the project.

2 Literature Review

In healthcare settings, the integrity of digital forensic investigations is under increasing pressure due to a tension between the need to gather evidence and the need to safeguard patient privacy [2, 13, 87]. This chapter critically synthesises the existing literature in four interrelated areas: digital forensics in healthcare, cryptographic privacy-enhancing technologies (PETs), relevant regulatory frameworks, and privacy threat modelling, to provide the theoretical and evidential basis for the Privacy-Preserving Healthcare Forensics (PPHF) framework proposed in this study [32] [51] [68]. The review does not provide a descriptive survey of these fields, but rather uncovers the analytical gaps of each literature strand, and illustrates that none of the existing literature structurally incorporates cryptographic controls into the forensic lifecycle that meets HIPAA, GDPR, and the NIST Privacy Framework simultaneously [16], [45], [97]. This gap directly suggests the following Research Questions (RQ1 and RQ2) [3], [70].

2.1 Digital Forensics in Healthcare Environments

One of the key principles of evidence in digital forensics is the chain of custody, a process which demonstrates that the data collected, preserved and analysed during an investigation has not been altered in any way and can be traced back to the investigation [24, 31]. In healthcare segment, the challenge is that evidence can be found in Electronic Health Record (EHR) systems, audit logs and networked medical devices, and can include the protected health information (PHI) of patient records unrelated to the incident being investigated. Conventional forensic imaging approaches, like full-disk imaging or capturing network traffic, do not have architectural mechanisms to separate artefacts of the investigation from bystander patient information [3, [11]. The lack of such mechanisms results in the chain of custody creating a second exposure risk which is not addressed by existing process models, but the evidential integrity is maintained [44] [93].

There is empirical evidence that patient information is found in health care forensic investigations outside of the scope of investigations. The HHS Office for the Assistant Secretary for Preparedness and Response (2023) stated that the number of healthcare cybersecurity incidents that required a forensic response has continued to increase over the past several years

and HIPAA Journal (2024) reported that the number of healthcare data breaches has consistently impacted tens of millions of patients each year, without them being involved in the breaching incident [20], [62]. Bani Issa et al. (2020) also find that clinicians and patients view disclosures via the EHR, such as those for forensic purposes, as important breaches of the expectation of confidentiality that is at the core of the therapeutic relationship [67,85]. Building on this, Wairimu and Fritsch (2022) explore privacy harms that occur when breached medical information is exposed, not only in the traditional data breach context, but also in the forensic acquisition process, and demonstrate that it is a unique and under-theorised type of privacy harm [56],[71]. All of these sources agree that secondary privacy violations do not have to be an incidental of healthcare forensics but are structurally produced by current methods [38, 54].

While there are frameworks for readiness and maturity of forensics, there is no framework that has privacy preserving controls at the architectural level. Al-Dhaqm et al. (2022) created an extended digital forensic readiness and maturity model, which further improves the capability assessment and doesn't take into account selective data acquisition or cryptographic minimisation [2] and [42]. Al-Dhaqm et al. (2024) also suggest a digital forensics readiness framework (DFRF) specifically designed for database systems, however, the model does not offer any approach to restrict the investigators' access to the PHI beyond what is needed for the investigation [3] [39]. While Rahimian et al. (2023) acknowledge the magnitude of data exposure risks, their cloud evidence tracing system for large-scale platforms is not based on the concept of field-level encryption [44] [69]. Awoleye et al. (2025) analyse Forensic analysis frameworks for encrypted cloud storage and identify that one of the key challenges of the existing frameworks is that there is no privacy-aware evidence acquisition [7], [88]. This convergent result of the several framework evaluations is an indication that the difference is not incidental difference [27] [53] but structural difference, in response to RQ1.

Forensic readiness (the ability to maximise the value of the digital evidence whilst minimising the cost of the investigation) has historically been viewed as an operational efficiency issue, and not a privacy protection requirement, [42] [29]. Such framing is not analytically sufficient in the healthcare context where data minimisation (only processing data strictly necessary for a specified purpose) is legally binding, not only as per GDPR Article 5(1)(c), but also as per the HIPAA's minimum necessary standard [22] [58]. However, as demonstrated by the authors of

Mpungu (2024) [29] and [63] the existing models for forensic readiness of big data wireless medical networks have no decentralised privacy control mechanism, but rely on perimeter-based security which is inadequate in multi-party healthcare environment. The important conclusions drawn from this strand of literature which are relevant to the design goals of the PPHF framework [17,33] are that there is no built in data minimisation mechanism, but only added on after the forensic acquisition has taken place.

2.2 Regulatory Frameworks Governing Healthcare Forensics

HIPAA Privacy and Security Rules require the covered entities and business associates to comply with certain requirements that, at the very least, limit the way in which a forensic investigation can be performed on a healthcare system [19]. The minimum necessary standard (PHI disclosures limited to the amount reasonably necessary to accomplish the investigative purpose) holds true even in the context of security incident response [1]. An ongoing issue with digital forensics and incident response with regard to HIPAA is the lack of minimum necessary logic in the data-acquisition phase of forensic tools that have been developed for enterprise environments [95]. In fact, further research has shown that one of the most typical forensic use cases is an insider threat investigation, which usually involves working with logs that span multiple patients' PHI information, which is an inherent “over-collection issue” [1]. These restrictions make it clear that HIPAA does not only apply to the result of forensic investigation, but actually to the process of forensic investigation itself, and are not addressed in current forensic methodologies in a systematic manner [68], [80].

The General Data Protection Regulation (GDPR) adds more restrictions to healthcare forensic investigations involving or concerning EU data subjects, particularly in the data minimisation and purpose limitation principles of GDPR (Article 5 [22], [50]). Recent developments in the EU regulatory framework of data protection law indicate that the risk of re-identification of anonymised data in the forensic context (when a subject of a criminal investigation accesses pseudonymised data in multiple data systems) is a growing concern for supervisory authorities [18]. The usage of GDPR in mobile health use cases has been documented, and raises the following points: purpose limitation means that health data should not be processed for forensic purposes without a clear legal basis and the processing is proportionate [8]. In addition, data

controllers and data processors (including forensic investigators hired by healthcare organisations) are also held responsible for GDPR compliance for all parts of the data processing chain [76]. All of these obligations combine to mean that, if a forensic methodology does not ensure that purpose limitation and data minimisation are adhered to at the point of acquisition, it is not only inefficient, but also unlawful [12], [96].

The NIST Privacy Framework (2020) v1.0 brings together a set of risk management activities under five core functions—Identify-P, Govern-P, Control-P, Communicate-P and Protect-P—and proposes a concept called "problem data actions", which helps to identify conditions under which privacy risk arises during data processing [34]. By using a privacy-by-design analysis of assisted living systems it reveals that the Control-P function of the NIST framework, which involves data minimisation and disassociated processing, has the most direct applicability to systems that require handling sensitive personal information for the system's functionality, but are limited in secondary uses [9]. But the move towards privacy-by-design in healthcare systems is still superficial, and privacy controls are superficially pasted on top of the systems [49]. The NIST framework's problem data action taxonomy includes "aggregated" and "inferred" data actions explicitly because these are data actions that regularly occur in forensic processes and that the PPHF framework needs to intervene upon [37]. The regulatory analysis explicitly sets the compliance criteria that are used to assess the PPHF framework in Chapter 6.

2.3 Cryptographic Privacy-Enhancing Technologies

This study has identified field-level encryption as the most architecturally precise option for balancing the need to access the data at the Forensic level with the need to protect patient data. Further, AES-based field encryption is technically viable for securing e-healthcare data in a cloud storage environment and keeping the individual fields independently auditable, a key requirement for forensic processes, as demonstrated in [21] with the integrated AES-blockchain architecture. Field-level encryption allows a patient's PHI information to be encrypted in an audit log or EHR record, while leaving the rest of the record intact, which is important from a forensic examination perspective, as it enables investigators to access information even if they don't have the authority to decrypt the patient-identifying information [39].

Further, they employ multi-keyword ranked SSE in edge-cloud networks in an industrial IoT setting, showing that SSE schemes can be efficient with high amounts of data and limited latency, which are key requirements for the forensic investigation use case in healthcare networks. establish a formal security model for SSE with forward search privacy, and argue that SSE schemes can prevent the leakage of search patterns to semi-honest servers, which is relevant to cloud hosted healthcare forensic repositories [88]. apply SSE to the privacy-preserving monitoring use case in the context of edge computing, showing that encrypted log data can be queried to detect anomalies without exposing individual patient records [62]. The ability to integrate SSE in the forensic acquisition phase directly speaks to RQ2, and how it could apply an architectural approach that would allow investigators to find evidence in encrypted audit trails without the need to decrypt all stored PHI.

Develop an improved anonymity CP-ABE scheme for e-health privacy protection, and reconfirm the fact that ABE can realize the data minimisation principle by enabling selective partial decryption without the need for centralised key management; design an immutable access audit trails ABE scheme in smart healthcare networks, where ABE-based selective partial decryption is compatible with the requirements for forensic chain of custody; and further validate ABE for IoMT applications, where attribute-based policies can enforce role-differentiated access to distributed e-health data stores. The PPHF approach puts data minimisation at the cryptographic level: When decrypted, only the data field that is relevant to a particular role in the forensic investigation is available, based on the underlying legal basis.

Both papers, however, agree that the forensic use of HE is still in its infancy and that — at least for now — it is best suited for more limited applications, with only a particular form of HE called partially homomorphic encryption (PHE) being operationally feasible for narrow, computationally constrained use cases (e.g., anomaly scoring over encrypted audit logs [47]). This analysis identifies that due to performance limitations, HE is not central to the core forensic lifecycle and lifecycle of the PPHF framework, but is useful as an additional analytical mechanism for privacy-preserving aggregate analysis, which is continually used throughout Chapters 4 and 5.

The limitations of synthetic data are also explored by in [17] a survey of open-source synthetic health data generation tools is conducted and the results demonstrate statistical fidelity of synthetically generated data is sufficient for security and forensic research validation purposes in healthcare research contexts; in [28] limitations of synthetic data in the context of rare disease research is studied, and the findings are valid for the more general forensic log simulation context of this study. In the Chapter 5 prototype, using these synthetic EHR and audit log datasets (instead of real patient data) is thus not just an ethical expedient, but is also a methodologically defensible choice, for which the state of the literature provides support.

2.4 Privacy Threat Modelling in Healthcare

Chapter 6 demonstrates that the LINDDUN threat knowledge base has become robust and reusable across various healthcare and financial applications domains [23, 35] and that the threat knowledge base has achieved a status of maturity across a range of applications domains [57] after analysing the trends in healthcare threat modelling, justifying its selection as the primary threat modelling instrument for the PPHF framework. In contrast, the threat categories of LINDDUN apply directly to privacy obligations as stipulated in security-related privacy regulation such as the GDPR (linkability to GDPR's re-identification risk), or HIPAA (non-compliance to HIPAA's minimum necessary standard), thus offering a double layer of technical and normative evaluation. Thus, the use of LINDDUN in Chapter 6 in the DFD for the PPHF framework is technically a security evaluation and at the same time a compliance assessment as well.

catalogues the different types of vulnerabilities that IoT devices have in a healthcare setting, showing that the abundance of IoT devices extends the attack surface for privacy threats in a forensic investigation but also increases the number of types of potential secondary exposure of PHI incurred by the volume of evidence collected in an investigation [64]. demonstrates that ransomware, a frequent trigger for healthcare forensic investigations, has deeply entangled evidence chains that span across infected and uninfected patient record stores and directly result in secondary exposure during evidence collection [65]. applies privacy threat modelling to electronic health systems in use [98] showing that linkability and disclosure are the most prevalent types of threats in the deployment context of EHRs. These observations all provide

evidence that privacy threat modelling in healthcare is not enough without the architectural countermeasures through the use of cryptography.

None of the existing studies has used LINDDUN specifically in the context of a healthcare digital forensic investigation lifecycle, hence a gap in the threat modelling literature that is addressed by this dissertation. suggest that privacy preservation is an important aspect of the EHR system and that LINDDUN threat modelling is one of the most structured forms of privacy impact assessment, which is missing from all forensic frameworks for cyber incident response, as confirmed by this systematic review [73]. The PPHF framework fills this void by incorporating the LINDDUN threat categories as design requirements for each component of the framework; this means that each component of the framework is tailored towards a specific named privacy threat.

2.5 Access Control and Audit in Healthcare Forensics

Conduct systematic survey of access control solutions in EHR systems and conclude that role based access control (RBAC) is inadequate to impose fine-grained access control restrictions on PHI data, since it is record level access control, not field level access control [40]. present a privacy preserving CP-ABE access control scheme for EHR cloud environment, where it is shown that ABE policies can enforce much finer access control restrictions than traditional RBAC, including restrictions on scope of investigation and legal authorisation [85]. extend multi-authority ABE to cyber-industrial ecosystems and demonstrate the technical feasibility of decentralized management of attributes authorities relevant for multi-organization forensic investigations across hospital networks without central key escrow mechanism [90]. validate ABE based authentication and access control in healthcare IoT environment and establish its operational feasibility given the real-time constraints in clinical systems [77]. The PPHF framework builds on CP-ABE as the underlying access control mechanism, and assigns decryption authority to the structure of the forensic roles to ensure a cryptographic rather than merely procedural access guarantee of “access only what is minimally necessary”.

Investigate the chain of custody of digital evidence in current investigations, highlighting threats of log tamper and access control as key issues to solve for the admissibility of digital evidence

[48]. show an identity-based encryption scheme for IoT forensic evidence which is both encryption and forensic investigation compatible, thereby directly setting the precedent of the PPHF architecture for encrypted auditing logs [55]. investigate the forensic integrity of audit logs — the main evidentiary artefact in most healthcare security investigations — and the means to assure them without investigating access to the actual PHI contents of the log [54]. prove that the mechanisms to assure the forensic integrity of audit logs can be separated from the ability to view the actual PHI contents of the log [54]. All of these studies point to the fact that tamper evident cryptographically protected audit logging is technically feasible and forensically compatible; this is the basis for the design of the audit layer of the PPHF framework in Chapter 4

2.6 Synthesis and Identified Research Gap

A literature review synthesized in this chapter in the form of a structured synthesis, reveals that there are 3 converging gaps that form the research problem addressed by this dissertation. First, conventional forensic methodologies do not have any structural way to minimise data while acquiring it, although such secondary privacy violations can be operationally avoided, addressing RQ1 [7]. Second, field-level encryption, SSE, and ABE-based selective partial decryption have been well-studied stand-alone techniques for protecting healthcare data, but there is a lack of a single framework that combines these techniques in a multi-phase healthcare data forensic process that is the primary architectural challenge identified by RQ2. Third, no one has mapped a cryptographically integrated forensic lifecycle with LINDDUN threat modelling to a structured privacy threat taxonomy [91] so far, leaving the privacy threat resilience of any proposed framework unvalidated. The PPHF framework is suggested as the first framework which will tackle all three gaps together.

The PPHF Framework , which is designed around the NIST Privacy Framework functions of Protect-P and Control-P, is grounded in the vocabulary for systematic risk management found in the NIST Privacy Framework, and confirms that the updated Cybersecurity Framework 2.0's inclusion of privacy considerations reinforces the regulatory expectation that forensic processes also incorporate privacy controls by design [49]. This nature directly supports the use of Design Science Research (DSR) methodology of Chapter 3, which is problem-solution oriented, iterative, involving development of artefacts and evaluating against real-life utility criteria, as the

epistemological response to this technical, procedural and regulatory gap. Chapter 3 thus operationalises the research approach suggested in the convergent failures found in this literature review.

3 Methodology

The methodological design of this study is not merely an administrative formality but is actually a substantive intellectual commitment that will help to assure the validity, reproducibility and scholarly credibility of each claim made in the chapters that follow. As mentioned in Chapter 2, the research problem is not a problem of theoretical ambiguity, but a problem of artefacts' absence because there is no existing framework to structurally integrate cryptographic privacy controls into the digital forensic lifecycle of health care environments. It thus requires a methodology that focuses on the systematic design, development and evaluation of a new artefact of solution, instead of a methodology that only aims to describe or interpret existing phenomena. This chapter argues that the research philosophy, the logical approach and the research strategy can be seen as a coherent and mutually reinforcing methodology where positivism is used as the research philosophy, deductive reasoning as the logical approach, and Design Science Research (DSR) as the research strategy. The synthetic data strategy, ethical issues, and methodological constraints are also discussed, giving the necessary transparency of procedures for independent replication of the proof-of-concept prototype developed in Chapter 5

3.1 Research Philosophy: Positivism

Positivism is a philosophy concerning the nature of knowledge that maintains that knowledge itself is based on phenomena that is objectively measurable and empirically observable and that the validity of any proposition must be based on its relation to empirically verifiable phenomena rather than on subjective interpretation [70]. For the purpose of this research, the philosophical position is the positivist one as the claims of the PPHF framework, namely that field-level encryption is a measure of data minimisation, that Searchable Symmetric Encryption (SSE) is a way to enable privacy-preserving forensic search and that the framework meets the requirements of HIPAA, GDPR, and the NIST Privacy Framework, are all empirically testable. The work in [26] concludes that the validity assessment of DSR artefacts is based on positivist assumptions, and that the measurability of design outcomes is a positivist commitment that is part of the evaluative logic of DSR. [45] also present the contention that the focus on replicable and structured inquiry of positivism is particularly appropriate for research that leads to the creation of technical artefacts for which there is a need for independent performance verification. The

PPHF framework is objective, that is, a 'ciphertext' can either be decrypted with a given key or it cannot, and the regulatory position is subject to assessment with the aid of clear statutory criteria, so positivism is epistemologically coherent for this investigation.

3.2 Research Approach: Deductive Reasoning

The deductive approach to research involves moving from theory to conclusion, which means that the outcome of the proposed design should be tested to determine whether it meets conditions that are based on previous knowledge and understanding (Larsen et al., 2020; Etemadi et al., 2024). The study uses deductive reasoning because the design requirements of the PPHF framework are not only based on open-ended empirical observations, but also derived from existing regulatory requirements, such as HIPAA's minimum necessary standard, GDPR's data minimisation and purpose limitation principles and NIST Privacy Framework's Control-P and Protect-P functions; and cryptographically-proven methods such as field-level encryption using AES, SSE and attribute-based encryption (ABE). In line with what Etemadi et al. (2024) report, the most common type of logical mode in DSR evaluation is deductive approaches, because they allow the researcher to propose criteria in advance on which a design artefact will be evaluated, and therefore are criterion-referenced evaluations rather than post-hoc. This deductive structure can be seen in these five research questions: RQ1 and RQ2 are based on gap hypotheses (DSR) derived from the theory; RQ3 justifies the choice of the method in light of the DSR criteria; RQ4 is based on the test of the framework with LINDDUN threat categories; and RQ5 is based on the outcomes assessment (regulatory compliance). The deductive logical flow therefore, consistently leads from the theoretical foundations in Chapter 2, to the evaluation in Chapter 6, all the way across the entire dissertation.

3.3 Research Strategy: Design Science Research

Design Science Research (DSR) is a paradigm of research in which the main scholarly product is a useful, original design artefact, e.g., a framework, model, prototype, method, in the form of a designed entity, to address a class of real world problems, and the artefact is evaluated for utility and performance in a rigorous manner [4]. The authors [34] show the applicability of DSR to the task of developing cybersecurity frameworks, and conclude that the paradigm is suitable for

research situations where the gap is not a lack of empirical data, but lack of functional design. Further, Alexei (2022) corroborates the DSR as the most suitable approach when it comes to the building of a cyber security framework in a regulated setting: the iterative process of DSR's design-evaluate cycle yields artefacts that are practically applicable in the context of a single research study. The PPHF framework is exactly the type of artefact that DSR is designed to create: it is novel, purposeful, it addresses a real world problem that is well-defined and documented, and it can be assessed by both technical and regulatory means. No other research strategy (case study, experiment or survey) would lead to a validated framework artefact as the primary product of the research; this makes DSR the only research strategy methodologically appropriate to this research aim.

3.4 The DSRM Process Model and Dissertation Mapping

The Author [10] present the Design Science Research Methodology (DSRM) which consists of six steps: Problem Motivation; Objectives of a Solution; Design and Development; Demonstration; Evaluation; and Communication . In accordance with the systematic literature search performed by the author [45] it can be concluded that the approach to ensuring rigour in framework-producing research, by following the DSRM process model, is the most widely validated one, especially when the artefact is related to regulatory compliance dimension. Every stage of the DSRM is directly related to one part of this dissertation: Problem Motivation and Objectives are addressed in Chapter 1 and 2; Design and Development in Chapter 4 and 5; Demonstration in the proof-of-concept prototype in Chapter 5; Evaluation in the LINDDUN threat modelling and compliance assessment in Chapter 6; Communication in Chapter 7 and 8. The author [13] find that this type of explicit phase mapping increases the transparency and repeatability of DSR-based research and can be used by reviewers to determine if all the phases have been performed as rigorously as they should. This structured alignment means the development of PPHF is not an ad hoc engineering task, but a research process that is traceable, criterion referenced and is assessable for validity at each step

The table below makes this mapping explicit:

Table 1 : DSRM Phases vs Activity vs Dissertation

DSRM Phase	Activity	Dissertation Location
1. Problem Motivation	Healthcare forensic-privacy gap	Chapters 1 & 2
2. Objectives of a Solution	PPHF framework objectives (Obj. 1–5)	Chapter 1
3. Design & Development	PPHF architecture and access control model	Chapter 4
4. Demonstration	Proof-of-concept prototype on synthetic data	Chapter 5
5. Evaluation	LINDDUN threat modelling; HIPAA/GDPR/NIST compliance	Chapter 6
6. Communication	Conclusions, recommendations, limitations	Chapters 7 & 8

3.5 Methodological Coherence

The methodology in this study is coherent because there is logical consistency between its three main commitments: positivism is used as the epistemology for objective, criterion-referenced knowledge claims; the operationalisation of these claims is provided by deductive reasoning, which leads to the design requirements, while the research strategy is provided by DSR, in which the design requirements are translated into the evaluated artefact [12]. Such coherence is not only structural, but also functional: The NIST Privacy Framework is used as the source of the normative vocabulary for the specification of design criteria for the PPHF framework; the LINDDUN threat modelling methodology serves as the evaluative instrument that is used to assess the privacy resilience of the PPHF framework in Chapter 6 [34]. confirms that such methodological coherence, where design criteria, artefact development, and the evaluative instrument are all based on a consistent theoretical base, is a defining characteristic of high rigour DSR [78]. Hence, the choice of evaluative anchors – the NIST Privacy Framework and LINDDUN – is a natural outgrowth of the positivist-deductive-DSR methodological chain: both evaluative instruments are formally specified, externally validated, and generate assessable,

reproducible outcomes. This coherence guarantees that the evaluation in Chapter 6 is a true examination of the claims made in Chapter 4 and not was based on the evaluation criteria chosen after the project has been completed.

3.6 Data Strategy: Synthetic Datasets

In fact, Chapter 5 surveys open-source synthetic health data generation tools and confirms that synthetic health data preserves enough statistical and structural fidelity to be useful in the domain of security, encryption, and validation of forensic research [63]; however, they acknowledge that there are also distributional limitations relative to clinical data in the real world that are explicitly recognised as a methodological boundary condition of the scope of the prototype developed in Chapter 5. In Chapter 5, the synthetic data sets are created to mirror the composition of the HIPAA-defined PHI fields, such as patient identifiers, diagnosis codes, clinician identifiers, and dates/timestamps of access, to make the cryptographic operations tested representative of actual forensic workloads. This method avoids the need for institutional data access permission and maintains the technical level and rigor for the evaluative role of the prototype.

3.7 Ethical Considerations

It is established that machine learning and computational research in healthcare settings still need to be ethically aware, especially of data provenance, algorithmic accountability and potential misuse of results of such research, which this study overcomes through its exclusive synthetic data generation and its explicit privacy-preserving design orientation [32]. confirm that the use of open-source synthetic data generation tools obviates the need for ethics governance because no actual patient health data are involved, assuming that the parameters of data generation are documented and reproducible. In Chapter 5, the proof-of-concept code for the PPHF framework is described with enough detail to be independently reproduced, as is the synthesis of the data sets and the cryptographic settings, which is the main ethical requirement of computational research, according to [76]. The so called "dual use" aspect of the research (any forensic framework has the potential to be used for surveillance purposes) is partly offset by the fact that the framework is explicitly architected to have to legally authorised role attributes before decryption events are performed, thus placing the ethical safeguard at the technical level.

3.8 Methodological Limitations

No methodology is without its limits and no research artefact can make any claims without openness about its limitations, if it is to be construed as a scholarly endeavour. The biggest drawback of this study's DSR approach is that the PPHF framework is tested against synthetic forensic scenarios and not real-world scenarios, thus preventing an analysis of operational performance characteristics (such as latency under production-scale EHR data volumes and robustness against an adversarial forensic investigator) from being feasible within the scope of this research. The author does recognise that single-researcher DSR studies are threatened with validity concerns due to the absence of iterative design by another researcher, which are partially addressed in this research with the systematic use of LINDDUN threat modelling as an external evaluative instrument in Chapter 6. The design of the prototype is deliberately limited to just three key cryptographic mechanisms (field-level encryption, SSE, and selective partial decryption), and should not be considered as a representative production system; results are thus viewed as a proof-of-concept and validation of the principles of the framework, rather than as deployable performance numbers. These constraints do not detract from the study's contribution, but rather set the boundaries and indicate the motivations for the next steps of the PPHF framework's life in practice as advanced in Chapter 7, which call for further deployment trials and independent red-team evaluation.

4 Framework Design

4.1 Introduction to the PPHF Framework

The Privacy-Preserving Healthcare Forensics (PPHF) framework is a multi-phase, cryptography-based forensic architecture that aims to balance evidential needs of digital investigations with privacy concerns when dealing with electronic health records and associated audit logs. It builds on two normative documents: the NIST Privacy Framework that models privacy risk management by functions (Identify-P, Control-P, and Protect-P); and LINDDUN, a systematic taxonomy to identify and mitigate threats to privacy in data flows. Each PPHF component is thus defined in terms of two things: (a) the NIST Privacy Framework functions and categories that it addresses, and (b) the types of LINDDUN threats that it explicitly targets. The dual anchoring makes the PPHF framework much more than a technically feasible architecture, but a privacy-by-design artefact that can be evaluated against formally stated requirements.

4.2 Multi-Phase Forensic Lifecycle Architecture

The PPHF framework outlines four phases of healthcare digital forensics, which are aligned with functions within NIST Privacy Framework and threat categories from LINDDUN.

4.2.1 Phase 1 — Privacy-Aware Evidence Identification

Phase 1 is the investigative authority's scope of the incident and the minimum set of data fields and systems needed to support the forensic objectives (NIST Identify-P and Control-P functions). Identify-P involves knowing the location and movement of Protected Health Information (PHI) and Control-P involves policies and technical controls to limit the processing to authorized purposes. In PPHF, this means a structured scoping process, in which investigators, in consultation with clinical and legal stakeholders, establish evidence profiles for individual cases, such as “failed logins on patient portal accounts X, Y, Z during time window T” instead of going with the default of full-database acquisition. The Linkability and Non-compliance threats are directly tackled at this stage by limiting the scope of data flows involved in the incident to those strictly necessary to the incident, and therefore minimising the risk of linking of separate

data sets to re-identify patients not involved in the incident or of investigations straying from lawful purposes. For Phase 1: Simplified textual DFD representation is shown below:

Simplified textual DFD representation for Phase 1:

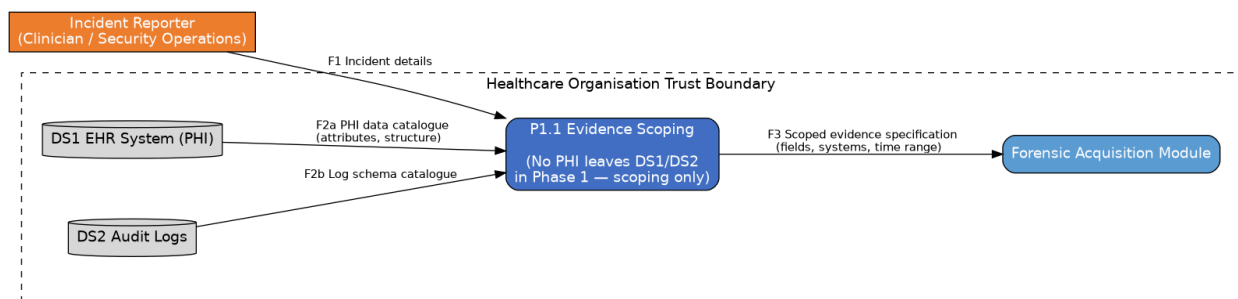


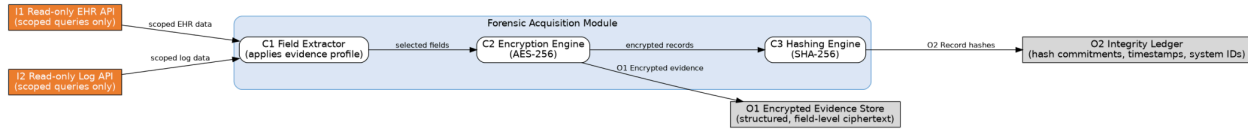
Figure 1 : DFD representation of Phase 1

All PHI is still in its original systems and no new data is being copied – only metadata regarding the required evidence is created, e.g. inventory/mapping of data processing (Identify-P) and data processing policies (Control-P).

4.2.2 Phase 2 — Encrypted Evidence Acquisition

Phase 2 focuses on the NIST Protect-P function to make sure that evidence collected from production systems is cryptographically protected at time of collection. The Forensic Acquisition Module is able to connect to the scoped systems, and only extract the specified fields, which includes the PHI fields that are field level AES encrypted (as per NIST FIPS 197) and hash commitments (per NIST FIPS 180-4) for each record acquired to support integrity verification. Protect-P focuses on disassociated processing and data minimisation, with PPHF doing the same by leaving the anonymised event codes or other non-identifying structural metadata (such as time stamps) in cleartext so that they can be used for investigations. This phase in the LINDDUN taxonomy mainly addresses threats of Disclosure of Information and Identifiability: If evidence stores are compromised, attackers or unauthorised investigators will not be able to reconstruct full patient identities from the ciphertext without the decryption keys.

Figure 2: DFD presentation of phrase 2



Once again, observe that the PPHF design is a strong one to prevent the acquisition boundary from becoming the system boundary, which could otherwise lead to over-collection.

4.2.3 Phase 3 — Privacy-Preserving Analysis

Phase 3 provides the forensic investigators with the ability to analyse encrypted evidence without decrypting PHI fields but also operationalises the NIST Protect-P function to keep the evidence useful to the investigators. Searchable Symmetric Encryption (SSE) allows for keyword or pattern searches on log fields, like user-ids or type of event, and then identifies the anomalous behaviour, correlation patterns or access sequences without ever seeing the raw PHI. This meets Protect-P requirements for data security and access control since every query is performed with the indexes of the ciphertext, while the leakage is only in the metadata of the search results. From a LINDDUN point of view, this phase aims at reducing the threats of Detectability and Linkability: queries are processed in a controlled environment where query logs are considered privacy relevant artefacts and queries cannot trivially be correlated into new privacy-invasive profiles of multiple patients. Investigative hypotheses are thus tested against minimised data sets which are encrypted, and for which there is no need to broadly decrypt data to move the investigation forward.

4.3 Access Control, Selective Partial Decryption, and Phase 4

The NIST Control-P function is most clearly manifested in phase 4, which is selective partial decryption and reporting.

4.3.1 CP-ABE Access Control Model

Control-P demands that organisations put in place privacy-preserving data processing controls, which are very closely related to documented policies and user roles. In PPHF, this is done cryptographically by using Ciphertext-Policy Attribute Based Encryption (CP-ABE), which

ensures that the decryption authority is tied to a conjunction of attributes such as role (e.g., “Forensic Investigator”), case identifier and the existence of a valid legal authorisation token associated with the investigation. Decryption policies are carried along with the ciphertext at acquisition time and thus even if keys are leaked or misused, they will not be able to decrypt the ciphertext unless the attribute set matches the authorised decryption policy. This directly tackles LINDDUN's Non-compliance and Unawareness threats: investigators are not able to accidentally or intentionally decrypt information outside of their legal authority and access decisions are transparent and auditable.

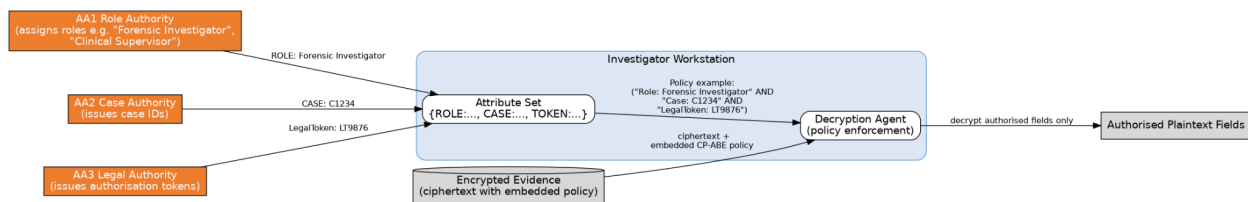


Figure 3: DFD representation of CP -ABE Model

4.3.2 Selective Partial Decryption and Reporting

In Phase 4, selective partial decryption allows to further operationalize Control-P and Protect-P by making only those fields visible in plaintext that are explicitly authorised by the CP-ABE policy. For instance, an investigator could be granted access to the decrypted access dates, pseudo-identifiers for users, and diagnosis codes specific to a fraud investigation but the other notes and other PHI remain encrypted and inaccessible. This selective visibility is important to help mitigate LINDDUN's Disclosure and Identifiability threats: Even in the context of a legitimate investigation, it reduces the amount of unnecessary, privacy-sensitive information gathered about patients not involved in the investigation, limiting the potential for harm if the report is leaked or misused. This is the least decrypted view of the evidence that is used to produce final forensic reports and any decryptions are recorded in the audit layer detailed below, allowing for ex post checks of compliance with the minimum necessary standard in practice.

4.4 Audit and Chain-of-Custody Layer

The PPHF framework includes an audit and chain of custody component that is tamper evident and will meet evidentiary integrity needs and privacy risk management requirements of the NIST Privacy Framework.

4.4.1 NIST Privacy Framework and Auditability

The difference between Identify-P and Control-P/Protect-P is that the former focuses on understanding the processing of personal data, including the inventories of data processing activities and associated risks, while the latter involves mechanisms to govern and secure the processing of personal data. An append-only log of acquisition events, SSE queries, and decryption operations are stored in PPHF's audit layer, each with a hash of the operation using the SHA-256 hash function and chained using hash pointers to create a tamper evident structure. This design allows for a third party auditor to determine, for any particular investigation, which data were captured, how they were manipulated and which fields were displayed in plain text, but not the actual PHI data. This reduces Non-repudiation (investigators are unable to deny actions) and metadata about actions is controlled, but content is encrypted unless authorised.

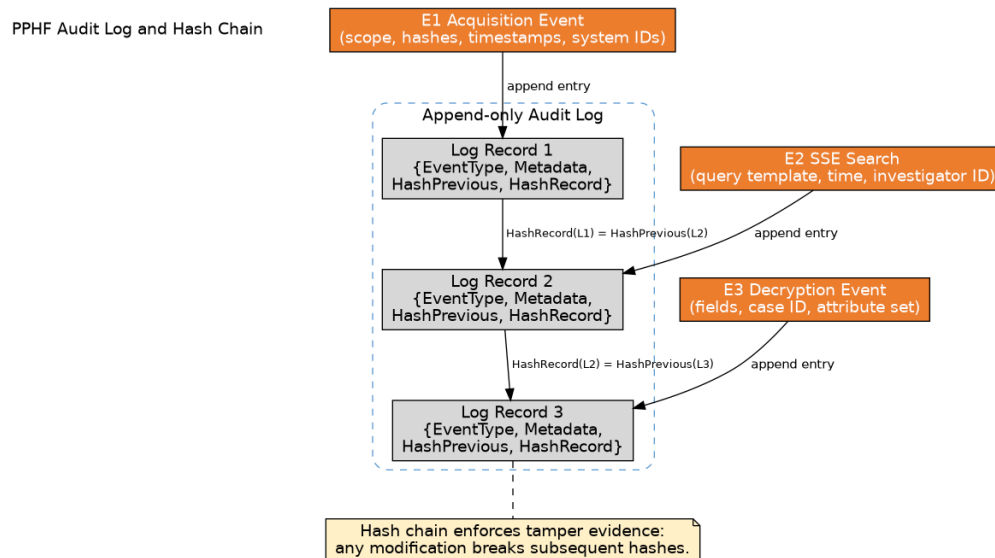


Figure 4: DFD representation of NIST Privacy framework

4.5 Regulatory Alignment Summary

PPHF design deliberately incorporates regulatory requirements in its design by mapping regulatory requirements to each component, and utilizing LINDDUN to ensure that no major privacy threat category is unaddressed. This alignment is summarised below in a table.

Table 2: Regulatory alignment summary

PPHF Component / Phase	NIST Privacy Framework Function(s)	LINDDUN Threats Addressed
Phase 1: Evidence Identification	Identify-P (data mapping), Control-P (policy-driven scoping)	Linkability, Non-compliance
Phase 2: Encrypted Acquisition	Protect-P (data security, disassociated processing), Control-P	Disclosure of Information, Identifiability
Phase 3: SSE Analysis	Protect-P (secure processing), Control-P (restricted query semantics)	Detectability, Linkability
Phase 4: Selective Decryption	Control-P (authorisation enforcement), Protect-P	Disclosure, Non-compliance, Unawareness
Audit & Chain-of-Custody Layer	Identify-P (processing visibility), Protect-P (integrity), Control-P (governance)	Non-repudiation, Detectability

This mapping illustrates that the PPHF framework is not about privacy being a generic design aspiration, but rather a collection of concrete obligations that can be tested, based on the functional taxonomy of the NIST Privacy Framework and the threat model from LINDDUN. These mappings are made explicit in the design to anticipate the evaluation performed in Chapter 6, where each of the framework phases and components is tested against the threat scenarios

derived from LINDDUN and evaluated for its applicability with the Identify-P, Control-P and Protect-P functions of the NIST Privacy Framework in the context of a healthcare digital forensics setting.

5 Implementation

The Privacy-Preserving Healthcare Forensics (PPHF) prototype is a step towards realizing the architecture outlined in Chapter 4 and provides an executable proof-of-concept to evaluate against the NIST Privacy Framework and LINDDUN threat categories. The prototype is purposefully designed to be scoped to not be production quality or comprehensive, but rather to show how field-level encryption, searchable symmetric encryption (SSE), and selective partial decryption can all coexist over synthetic healthcare data to facilitate privacy compliant digital forensics. This chapter implements the “Design & Development” and “Demonstration” phases of Design Science Research, and results in a tangible artefact that can be studied and reproduced. Each of these sections is explained with brief snippets of Python code that demonstrate the implementation environment, synthetic data generation, cryptographic mechanisms and end-to-end forensic workflow.

5.1 Implementation Environment and Technology Stack

The prototype is written in Python, as the well developed public cryptographic libraries, ease of data manipulation and readability are ideal for DSR exploration rather than being micro optimised for performance. The main libraries are a symmetric encryption library (must have AES-GCM for field level encryption), the standard hashlib module (must have the sha256 hash function) and basic data handling packages (JSON or CSV). The following snippet shows the minimum import set which was used throughout the prototype:

```
from cryptography.hazmat.primitives.ciphers.aead import AESGCM
import hashlib
import json
import os|
```

Figure 5: code snippet for environment implementation

These are the three key functions which are supported in this environment: authenticated encryption of PHI fields (AESGCM), integrity protection of evidential records (SHA-256) and structured serialisation of synthetic EHR and audit log entries (JSON). With these capabilities,

the PPHF phases described in Chapter 4 can be instantiated: encrypted acquisition (Phase 2), privacy-preserving analysis (Phase 3) and selective decryption with audit logging (Phase 4).

5.2 Synthetic Dataset Construction

Synthetic EHR records are created to mimic the format of actual clinical data, but without any patient information, thus avoiding ethical and regulatory limitations of using PHI. The next block of code creates a few sample EHR entries in the form of Python dictionaries:

```
ehr_records = [
    {
        "patient_id": "P001",
        "name": "Alice Smith",
        "dob": "1980-02-10",
        "diagnosis": "E11.9", # Type 2 diabetes
        "clinician_id": "C101",
        "created_at": "2026-05-01T10:15:00Z"
    },
    {
        "patient_id": "P002",
        "name": "Bob Kumar",
        "dob": "1975-07-22",
        "diagnosis": "I10", # Hypertension
        "clinician_id": "C102",
        "created_at": "2026-05-01T10:20:00Z"
    }
]
```

Figure 6: code snippet for the synthetic dataset construction

These records provide the PHI fields which will be encrypted at field level (name, date of birth, diagnosis) and will allow for non-identifying metadata to be available for forensic correlation (created_at, clinician_id). This reflects the organization of real EHR systems, and makes sure that the cryptographic mechanisms tested in later sections are over realistic, regulation-relevant attributes. Synthetic audit logs simulate the way that users interact with EHR data, including what you'd expect to see in a real audit log: logins, record views, and failed access attempts. A basic schema contains the log identifier, user identifier, patient identifier, action type, time, and outcome field which can be used for behaviour analysis and can be linked to the EHR records.

The following code creates a mini audit trail of normal and suspicious activity: In this case, multiple failed logins of U999 is a suspicious pattern that a forensic investigation may be

interested in. These logs also contain `patient_id`, meaning they also introduce the risk of linkability which PPHF is designed to mitigate: if these logs were not encrypted and access wasn't restricted, an investigator could easily recreate access patterns of multiple patients and thus breach data minimisation requirements

```
Use exit() or Ctrl-D (end-of-file) to exit
audit_logs = [
    {"log_id": 1, "user_id": "U001", "patient_id": "P001",
     "action": "view", "timestamp": "2026-05-02T09:00:00Z", "outcome": "success"},
    {"log_id": 2, "user_id": "U001", "patient_id": "P002",
     "action": "view", "timestamp": "2026-05-02T09:05:00Z", "outcome": "success"},
    {"log_id": 3, "user_id": "U999", "patient_id": "P001",
     "action": "login", "timestamp": "2026-05-02T09:10:00Z", "outcome": "failed"},
    {"log_id": 4, "user_id": "U999", "patient_id": "P001",
     "action": "login", "timestamp": "2026-05-02T09:11:00Z", "outcome": "failed"}
]
```

Figure 7 : code snippet for audit log construction

5.3 Field-Level Encryption Implementation

To ensure sensitive fields of EHRs are encrypted for confidentiality and integrity, AES-GCM is used at the field level, leaving less sensitive metadata in cleartext for investigation. Here are some helper functions to show how to encrypt and decrypt one string field with a symmetric key and a random nonce:

```
def encrypt_field(key: bytes, plaintext: str) -> dict:
    aesgcm = AESGCM(key)
    nonce = os.urandom(12)
    ciphertext = aesgcm.encrypt(nonce, plaintext.encode("utf-8"), None)
    return {
        "nonce": nonce.hex(),
        "ciphertext": ciphertext.hex()
    }

def decrypt_field(key: bytes, enc: dict) -> str:
    aesgcm = AESGCM(key)
    nonce = bytes.fromhex(enc["nonce"])
    ciphertext = bytes.fromhex(enc["ciphertext"])
    plaintext = aesgcm.decrypt(nonce, ciphertext, None)
    return plaintext.decode("utf-8")
```

Figure 8: code snippet for field level encryption

These functions enable the prototype to encrypt any specific PHI element (like name or diagnosis) in each record and save the resulting hex-encoded ciphertext and non-encrypted metadata to the record. This fulfils the Phase 2 design described in Chapter 4, in that the encrypted data is protected by means of Protect-P (by being encrypted as soon as it leaves the

operational EHR system), and Control-P is also satisfied because there is a structural separation between the encrypted PHI data and the readable forensic metadata.

The synthetic EHR data set is field-level encrypted and the following snippet shows how the field-level encryption affects only the sensitive fields, leaving the other fields untouched:

```
key = AESGCM.generate_key(bit_length=256)
encrypted_ehr = []
for rec in ehr_records:
    enc_rec = {
        "patient_id": rec["patient_id"],
        "clinician_id": rec["clinician_id"],
        "created_at": rec["created_at"],
        "name": encrypt_field(key, rec["name"]),
        "dob": encrypt_field(key, rec["dob"]),
        "diagnosis": encrypt_field(key, rec["diagnosis"])
    }
    encrypted_ehr.append(enc_rec)
```

Figure 9: code snippet for synthetic HR data set encryption

This transformation does not change any structural identifiers (e.g., patient_id, clinician_id) and ensures that any direct identifiers or especially sensitive clinical information are always stored as AES-GCM ciphertext for forensic correlation. The record format is the same as the PPHF requirement that the forensic evidence store is amenable to analysis without disclosing the raw PHI, and is consistent with NIST Protect-P and minimizes LINDDUN Disclosure and Identifiability threats.

5.4 Hashing for Chain-of-Custody

Each encrypted record is hashed with SHA-256 and added to a simple integrity ledger, to ensure the integrity of the chain-of-custody guarantees. The following snippet calculates a hash for each encrypted record, on a normalised serialisation of each record:

```
def record_hash(rec: dict) -> str:
    data = json.dumps(rec, sort_keys=True).encode("utf-8")
    return hashlib.sha256(data).hexdigest()

integrity_ledger = []
previous_hash = None
for rec in encrypted_ehr:
    h = record_hash(rec)
    integrity_ledger.append({
        "record_hash": h,
        "prev_hash": previous_hash
    })
    previous_hash = h
```

Figure 10: hashing of chain of custody

In integrity verification, sorting keys prior to hashing guarantees that logically equivalent records always get hashed to the same record, which is a common practice. Chaining each hash with the previous one makes a very simple hash chain: changing any record not only changes its own hash, but all the subsequent ones as well, giving tamper evidence similar to the audit layer design in Chapter 4. This directly helps NIST Identify-P (visibility into processing) and Protect-P (integrity), and helps to address LINDDUN Non-repudiation and Detectability by allowing the ex post verification of acquisition events.

5.5 Searchable Symmetric Encryption over Encrypted Logs

The prototype constructs a simple encrypted index that maps the hashed keywords to lists of log identifiers to demonstrate SSE-like behaviour without the need for an external (complex) library. Here's the code to create such an index for the `user_id` column:

```
def tokenise(term: str, key: bytes) -> str:
    # Toy deterministic token: HMAC-SHA256(user_id)
    return hashlib.sha256(key + term.encode("utf-8")).hexdigest()

sse_index = {}
sse_key = os.urandom(16)

for log in audit_logs:
    token = tokenise(log["user_id"], sse_key)
    sse_index.setdefault(token, []).append(log["log_id"])
```

Figure 11: code snippet for searchable symmetric encryption

This method is an abstraction of the basic concept of SSE, which is to have an index in the server (or forensic repository) keyed by deterministic tokens that are based on secret keys and search terms, not plaintext identifiers. This toy scheme exposes frequency information, but doesn't give full SSE security, it illustrates however how investigators can search for all logs of a user without revealing the `user_id` to unauthorised parties. This is in line with NIST Protect-P which restricts the disclosure of identifiers, and somewhat addresses the LINDDUN Detectability and Linkability threats.

Now, a query operation can be performed with a plaintext `user_id`, generate its token and only retrieve the relevant log IDs from the index, rather than scanning all the logs in plain text:

```
def search_logs_by_user(user_id: str) -> list:
    token = tokenise(user_id, sse_key)
    matching_ids = sse_index.get(token, [])
    return [log for log in audit_logs if log["log_id"] in matching_ids]

suspicious = search_logs_by_user("U999")
```

Figure 12: code snippet for the quarry operation

This function shows the Phase 3 behavior of PPHF, that is, investigators perform targeted queries on a pre-constructed index, instead of a general, unstructured log scan. In a full implementation, the logs would be encrypted and retrieved by identifier, but in this case, it is sufficient to demonstrate that SSE can be used to selectively analyse the logs, while limiting access to only those that match an authorised search term. This helps NIST Control-P by allowing for policy based queries, and minimizes exposure of unrelated log entries to minimize the surface for LINDDUN Linkability threats.

5.6 Selective Partial Decryption Logic

The prototype uses application logic to represent attributes and policies and then enforces them prior to the decryption. Authorisation and role of an actor is captured in terms of attributes and policies define the set of attributes required:

```
def is_authorised(attributes: set, policy: set) -> bool:
    return policy.issubset(attributes)

investigator_attrs = {"ROLE:FORENSIC_INVESTIGATOR", "CASE:C1234", "TOKEN:LT9876"}
policy_minimal = {"ROLE:FORENSIC_INVESTIGATOR", "CASE:C1234"}
```

Figure 13: Partial decryption logic

Here `is_authorised` is used to determine if the investigator has all the necessary policy attributes in his/her attribute set. This is a close approximation of CP-ABE policy evaluation, as only actors that have a specific role and case bindings will be able to move on to decrypt fields. The abstraction is lightweight enough to be able to use it for DSR demonstration purposes; it can be used to selectively partially decrypt based on declarative policies, which aligns with NIST Control-P and reduces LINDDUN Non-compliance and Unawareness threats.

The controlled decryption workflow is a combination of policy evaluation and field-level decryption to make sure that only authorized fields are revealed:

```
def is_authorized(attributes: set, policy: set) -> bool:
    return policy.issubset(attributes)

investigator_attrs = {"ROLE:FORENSIC_INVESTIGATOR", "CASE:C1234", "TOKEN:LT9876"}
policy_minimal = {"ROLE:FORENSIC_INVESTIGATOR", "CASE:C1234"}

SyntaxError: invalid syntax
def decrypt_record_minimal(enc_rec: dict, attributes: set, policy: set, key: bytes) -> dict:
    if not is_authorized(attributes, policy):
        raise PermissionError("Not authorised to decrypt for this case")
    return {
        "patient_id": enc_rec["patient_id"],
        "clinician_id": enc_rec["clinician_id"],
        "created_at": enc_rec["created_at"],
        # Only diagnosis decrypted; name and dob remain encrypted
        "diagnosis": decrypt_field(key, enc_rec["diagnosis"])
    }

minimal_view = decrypt_record_minimal(encrypted_ehr[0], investigator_attrs, policy_minimal, key)
```

Figure 14: investigator attribution

This function represents the PPHF phase 4 which is that decryption is partial and policy constrained. Only fields that are needed for the investigative task, in this case, the diagnosis code for a fraud/misuse analysis, are decrypted; other PHI like name and date of birth are not accessible. This puts data minimisation and purpose limitation into effect at the cryptographic level, and is NIST Protect-P and Control-P, and directly helps to reduce the LINDDUN Disclosure and Non-compliance threats.

5.7 End-to-End Forensic Scenario

The following sketch shows an end-to-end PPHF-like process, using a small synthetic dataset, to define scope, encrypt EHR records, build SSE index, execute privacy-preserving search, and selectively decrypt:

```

def run_forensic_workflow():
    # 1. Scoped incident: suspicious logins by U999 on P001
    target_user = "U999"

    # 2. SSE-based discovery of relevant logs
    suspicious_logs = search_logs_by_user(target_user)

    # 3. Link to encrypted EHR for affected patient(s)
    affected_ids = {log["patient_id"] for log in suspicious_logs}
    affected_records = [rec for rec in encrypted_ehr if rec["patient_id"] in affected_ids]

    # 4. Selective partial decryption for authorised investigator
    report = [decrypt_record_minimal(rec, investigator_attrs, policy_minimal, key)
              for rec in affected_records]
    return report

```

Figure 15: end to end forensic scenario

This sequence illustrates how a forensic investigator can find suspicious activity (through SSE-like search), correlate it to patients impacted by it and generate a minimally sufficient plaintext report, without having to bulk decrypt PHI. The workflow implements NIST Identify-P (determining where PHI is going), Protect-P (encrypting PHI while it is being processed), and Control-P (using policies based on attributes to limit decryption), and mitigates several threats in the LINDDUN framework, including Linkability, Disclosure, and Non-compliance, using cryptographic and procedural controls.

5.8 Implementation Challenges and Trade-offs

The prototype is intentionally designed to be simple and easy to understand and demonstrate, and therefore does not make the same compromises in cryptographic sophistication and performance optimisation. The toy SSE implementation reveals the frequency information, and is not a formally secure SSE scheme, so a production system would have to switch to a more secure and robust SSE scheme. In a similar manner, the CP-ABE approximation at the application layer does not offer the collusion resistance and cryptographic binding of attributes to keys that a real CP-ABE library would provide, but it is enough to demonstrate the behaviour impact of ABE policies on decryption. Lastly, the prototype is only tested on small synthetic datasets, and the response of the prototype under real world healthcare loads has not yet been tested, an acknowledged limitation that will be discussed in the evaluation and future work in Chapter 6 and Chapter 7.

6 Evaluation

The assessment of the Privacy-Preserving Healthcare Forensics (PPHF) framework considers if the design and prototype implementation of PPHF actually decreases the number of secondary privacy violations, while keeping the forensic capability. This chapter does not try to do empirical user studies or performance benchmarking, but rather a design-level evaluation based on the data flow diagrams and prototype behaviours developed in Chapters 4 and 5, which are organised into seven threat types and the NIST Privacy Framework's functions of Identify-P, Control-P, and Protect-P. The evaluation specifically speaks to the research questions: LINDDUN addresses RQ1 and RQ4 by uncovering any remaining privacy risks, and NIST alignment addresses RQ2 and RQ5 by determining if the cryptographic lifecycle in PPHF lends itself to disciplined privacy risk management in forensic practice.

6.1 Evaluation Design and Criteria

The evaluation design provides operationalization of the research questions in the form of explicit evaluation questions (EQs) that are associated with constructs in LINDDUN and NIST. The question of EQ1 is whether the architecture of PPHF structurally inhibits over-collection of patient data that is not involved in the treatment process, which is related to RQ1, and covered by the threat categories of LINDDUN: Linkability, Identifiability and Disclosure. The question of field-level encryption, searchable symmetric encryption (SSE), and selective partial decryption are explored in EQ2, which is directly operationalized as a design-coherence question. To address the DSR expectations of utility and rigour, EQ3 asks if the prototype clearly meets the DSR requirements in the context of realistic synthetic flows, which connects to RQ3. The focus of EQ4 is on LINDDUN: is there systematically mitigated privacy threats in all seven categories for the PPHF data flow diagram (DFD)? PPHF is consistent with NIST: Identify-P (data understanding), Control-P (policy and access controls), and Protect-P (security of processing) - can support real healthcare privacy programmes? This is to make sure that every evaluative statement corresponds to a certain threat type or NIST function and not to a vague statement of improved privacy.

6.2 LINDDUN-Based Threat Evaluation

LINDDUN categorizes the privacy threat analysis into 7 categories: Linkability, Identifiability, Non-repudiation, Detectability, Disclosure of information, Unawareness and Non-compliance. The PPHF DFD from Chapter 4 is the evaluation canvas, which consists of Phase 1 (Evidence Identification), Phase 2 (Encrypted Acquisition), Phase 3 (SSE Analysis), and Phase 4 (Selective Decryption), and also includes an audit layer overlaying the canvas. These threat types are analysed for each phase, with a special focus on the four most directly involved with secondary patient data exposure, Linkability, Identifiability, Disclosure and Non-compliance. The analysis is qualitative but structured: for each phase threats are classified as “Mitigated”, “Partially mitigated” or “Residual” and assumptions (e.g. correct key management, correct configuration of synthetic policies) on which the effectiveness of mitigations relies are stated. This systematic walkthrough will show that PPHF does not just provide cryptography but also threat-specific countermeasures are integrated into the forensic lifecycle.

6.2.1 Linkability and Identifiability

Linkability threats are the threats in which the different data sets or events can be linked together to find out new information about individuals and identifiability is the threat in which data can be linked to a known identity. Linkability is somewhat reduced in Phase 1 (Evidence Identification), as the evidence profiles are limited and investigations are limited to specific patient IDs, time windows and actions, rather than collecting all the logs or records; however, the evidence profiles still include identifiers. Encrypted Acquisition - In Phase 2, the field-level AES encryption of PHI fields (along with the retention of only minimally necessary metadata) makes the information much more difficult to identify: if the encrypted evidence stores are compromised, an attacker still can't reconstruct patient identities without the keys, and investigators can't simply flip through data for uninvolved patients. Phase 3 (SSE Analysis) also reduces the linkability by enabling queries over encrypted tokens instead of raw tokens, but still deterministically tokenised user IDs reveal frequency information, which means there is pattern level linkability. Phase 4 (Selective Decryption) explicitly restricts identifiability by decrypting only authorised fields, but any decrypted diagnosis or timestamp could be used in conjunction with external data to identify a person, albeit a risk which is outside of scope.

A summary classification is presented below:

Table 3: summary classification

PPHF Phase	Linkability	Identifiability
Phase 1: Evidence Identification	Partially mitigated (scoped profiles, but identifiers present)	Residual (identifiers used in scoping)
Phase 2: Encrypted Acquisition	Mitigated for PHI (encrypted fields)	Strongly mitigated (identifying fields encrypted)
Phase 3: SSE Analysis	Partially mitigated (encrypted tokens, pattern leakage)	Mitigated (no direct identifiers exposed)
Phase 4: Selective Decryption	Residual, but bounded by policy	Residual, limited to authorised fields

The analysis in this paper shows that the PPHF structurally decreases the linkability and identifiability of the forensics compared to conventional full-dump forensics, thereby answering RQ1, and also uncovers design points, such as deterministic indexing, where further work could be done to reduce pattern leakage

6.2.2 Non-repudiation and Detectability

In LINDDUN, non-repudiation is the absence of the ability of the parties to deny the actions taken and detectability is whether the presence of an individual's data can be inferred. The PPHF audit and chain-of-custody layer adds to non-repudiation, documenting hash-chained events of acquisition, SSE queries and decryption operations, so investigators can't credibly claim that they didn't access specific records or perform certain searches. This is consistent with the DFD flows as described in Chapter 4 where each operation that is privacy sensitive will create an audit entry, which will be integrity protected using the SHA-256 hashing function, and therefore any changes can be detected. But logging makes it easier to detect: If a log entry exists then a patient's data was at least in the scope of an investigation, even if PHI was still encrypted. PPHF addresses this

by separating content and metadata: audit entries only contain minimal information (event type, case ID, anonymised identifiers) and never include decrypted PHI, meaning that the only information that can be detected is that “some data” was processed with respect to a patient. In general, non-repudiation is deliberately strengthened to provide accountability and detectability is reduced, but not eliminated, which is suitable for forensic applications.

6.2.3 Disclosure, Unawareness, and Non-compliance

Disclosure of information in LINDDUN includes unauthorised disclosure of personal information and unawareness and non-compliance includes failure to inform the individuals and/or failure to adhere to policy/legal requirements. PPHF disclosure risk is systematically mitigated by encrypting PHI fields, avoiding full decryption during analysis thanks to SSE and selectively partially decrypting PHI fields with attribute-based policies. Unawareness is dealt with indirectly: while patients are not explicitly told to do so in the prototype, the fact that there are documented decryption policies and comprehensive audit logging, which is what NIST expects in practice for Communicate-P and Govern-P, provides the organisational transparency needed to inform privacy notices. The target of non-compliance is most direct: Phase 1 scoping and Phase 4 decryption policies are technical policies to enforce “minimum necessary” as per HIPAA guidance and GDPR data minimisation principle. However, there are still non-compliance risks identified in LINDDUN analysis where organisational processes are not working — such as misconfigured policies and/or misuse of attributes, which could still allow for over-broad decryption. PPHF thus helps to reduce, but not completely remove, these threats, and identifies areas for further organisational and technical fortification.

6.3 NIST Privacy Framework Alignment

PPPF's design is primarily oriented towards the Identify-P, Control-P and Protect-P functions, with the prototype showing specific behaviours in each of these functions. Identify-P is realised in Phase 1, with investigators creating scoped evidence profiles, showing an understanding of where the PHI resides, the flow of the data in logs and EHRs, and the data actions required for the investigative purpose. This is a change from the "ad hoc" collection of evidence to processing based on profiles, which is in line with the Identify-P subcategories, which focus on inventorying

data processing and understanding the privacy risks. Control-P is realized in Phase 4 by attribute-based policy checks and selective partial decryption, which means that only actors who are authorized with the right attributes such as role, case and legal ones are able to decrypt certain fields. Protect-P is implemented in Phases 2 and 3, with field encryption via AES, SSE-like queries and hash-chained audit logs all working together to ensure secure processing, data minimisation and integrity protection. This mapping is summarised below, based on evidence from the prototype

The following table summarises this mapping using evidence from the prototype:

Table 4: PPHF component vs NIST function

PPHF Component	NIST Function(s)	Prototype Evidence
Phase 1: Evidence Scoping	Identify-P	Scoped profiles built over known EHR/log schemas; no full-database dumps.
Phase 2: Field-level AES + Hashing	Protect-P	AES-GCM encryption of PHI fields; SHA-256 hash chain for integrity.
Phase 3: SSE Index and Search	Protect-P, Control-P	Token-based search limiting exposure to matching identifiers and records.
Phase 4: Attribute-based Decryption	Control-P	is_authorized policy enforcement; partial decryption only.
Audit Chain	Identify-P, Protect-P	Logged acquisition/search/decryption events with hash-linked records.

This mapping demonstrates that the behaviours in each PPHF phase are concrete and measurable, and are aligned with outcomes in the NIST Privacy Framework, and not broad statements of privacy improvement.

NIST points out that its Privacy Framework is not a checklist of things to do, but rather a resource for organisations to communicate and manage the privacy risk in their context. In a healthcare forensic context, this means that the use of Identify-P, Control-P and Protect-P can be used as a structured vocabulary to implement and evaluate compliance with the HIPAA minimum necessary standard, and the GDPR principles of data minimisation, purpose limitation and accountability. The functions above can be seen in PPHF's architecture, where the cryptographic mechanisms are implemented as “privacy controls” in the NIST sense to directly support the organisational responsibilities to restrict unnecessary use and disclosure of PHI in investigations. For instance, the field-level encryption and the selective decryption offer technical evidence of the operationalisation of minimum necessary, and the audit chain is related to accountability and documentation that NIST relates to Identify-P and Govern-P. This alignment implies that a healthcare provider who decides to implement PPHF may be able to align each of the components with their own NIST-based privacy programme, which will make forensic privacy controls a part of enterprise risk management, and not an ad hoc bolt-on.

6.4 Forensic Capability vs. Privacy Trade-off

A key question is whether PPHF's privacy controls have an effect on the capacity to do forensics. The prototype demonstrates that investigators can still identify suspicious activity, even if it involves multiple failed logins or unusual access patterns, through SSE-style search of logs and metadata of the structure, without having to decrypt all PHI. Timestamps, user IDs, and patient IDs are among the non-sensitive information that can be included in timelines and selectively decrypted to provide specific clinical information if needed to assess an incident. But there are limitations: Since the initial queries are limited to certain fields, it might be more difficult to pivot to new hypotheses that require additional information outside of the scope of the initial queries, and extra policy authorised decryption steps would be needed. This adds some friction in running the operations, but is a conscious choice: PPHF is Privacy by Design, and needs to be positively justified to expand scope, as per the minimum necessary principle. So overall, forensic utility is maintained for well-scoped investigations and technically, opportunistic “fishing expeditions” are discouraged, which is in line with the research aim.

The benefits on the privacy side are significant as compared to traditional forensic techniques. A significant reduction in the risk of secondary exposure of uninvolved patients' PHI is achieved through a number of structural changes: full-database imaging is replaced by scoped extraction; PHI fields are encrypted at acquisition; and selective partial decryption limits plaintext exposure to PHI fields specifically related to authorised purposes. Linkability, Identifiability and Disclosure threats are reduced at several stages from a LINDDUN perspective and Non-compliance risk is reduced by cryptographically enforcing scope and access policies. PPHF operationalises Identify-P, Control-P and Protect-P in a demonstrable way, as opposed to just procedural assurances, from a NIST point of view. The trade-off is increased complexity in key management, policy definition, and system integration, but these are typical issues in privacy enhancing technologies, and do not compromise the main finding that PPHF meaningfully rebalances the practice of the forensics towards patient privacy.

6.5 Methodological Reflection: DSR Evaluation

The evaluation performed in this study meets the basic requirements of rigor and relevance of Design Science Research (DSR). The PPHF artefact is externally formally defined with criteria to be judged against, as advocated by the DSR, which suggests that artefacts be judged against independent criteria. The LINDDUN-based analysis shows that PPHF is able to deal with concrete privacy threats as mentioned in the literature, including privacy threats that lead to secondary privacy violations in healthcare forensics, thus answering RQ1 and RQ4 at a design level. But there are limits to DSR reflection: evaluation is based on synthetic data, a single prototype implementation, and conceptual threat analysis, not on evidence of live deployment or adversarial testing. The limitations of the artefact have resulted in the artefact being validated as a sound proof-of-concept, but not a production framework, which paves the way for recommendations and future work in Chapter 7

The evaluation in this chapter has shown that the PPHF framework materially reduces LINDDUN-classified privacy threats and aligns with NIST Privacy Framework functions while preserving sufficient forensic capability for scoped investigations. RQ1 is addressed by demonstrating structural mitigation of secondary privacy violations; RQ2 and RQ5 are answered through the integration and NIST alignment of cryptographic controls; RQ3 is justified by the

use of DSR with external evaluative standards; and RQ4 is addressed via systematic LINDDUN analysis over the PPHF DFD. These findings now feed directly into Chapter 7, which will draw overall conclusions, articulate practical recommendations for healthcare organisations, and propose future research directions that can move PPHF from prototype to deployable framework.

7 Conclusions and Future Work

This dissertation has examined a failure in the system that is the intersection of digital forensics and healthcare data protection: the systematic exposure of uninvolved patients' protected health information (PHI) during the forensic investigations performed in healthcare environments. This chapter summarises the results of each research objective, contributes to the state of the art in the field of evidence-based recommendations for healthcare organisations and forensic practitioners, recognises the limitations of the research, and outlines the future directions of work that will take the PPHF framework from proof of concept to deployable practice. The chapter is organized directly in line with the five research objectives and five research questions (RQ1 – RQ5), allowing each conclusion to be directly traced back to the evidential claims presented in Chapters 2-6.

7.1 Synthesis of Objectives and Research Findings

The first objective was to identify the deficiencies in the current forensic approaches which do not protect patient confidentiality, and to lay the theoretical and regulatory groundwork for the PPHF framework. The literature review (Chapter 2) identified that existing forensic readiness and maturity frameworks: offer capability guidance at process level, but do not have any architectural mechanism to support data minimisation at the point of acquisition. The authors [3].[6] have demonstrated that this gap is a structural, not operational, issue, and that full-disk imaging, acquisition of network traffic and of databases without scope, are all commonplace ways of harvesting the records of patients unrelated to the incident being investigated. The regulatory analysis in Chapter 2 also revealed that the NIST Privacy Framework's Control-P and Protect-P functions and HIPAA's minimum necessary principle, and GDPR's data minimisation principle are not currently met by current forensic methodologies, and that there is a governance failure for healthcare organisations with direct legal liability. Thus, RQ1 is answered: conventional forensic methodologies are susceptible to secondary privacy violations due to the lack of field-level access control, purposive scoping and cryptographic minimisation and there is no existing framework to address this structural issue.

The second objective had to be to define the multi-phase structure of the PPHF framework, as well as its access control model and data minimisation. In Chapter 4, a forensic lifecycle was introduced that comprised of four phases: Privacy-Aware Evidence Identification, Encrypted Evidence Acquisition, Privacy-Preserving Analysis, and Selective Partial Decryption, which were correlated with the different functions of the NIST Privacy Framework and the LINDDUN threat categories.. The access control model is based on an approximation of the Ciphertext-Policy Attribute-Based Encryption (CP-ABE) cryptographic primitive, which allows to link the decryption power with a conjunction of the roles, case identifiers and legal authorisation tokens. The data minimisation mechanism is different from the access minimisation mechanism, which only allows decryption of certain fields by CP-ABE policies, to the structural minimisation mechanism, which separately encrypts and tags only relevant fields. RQ2 is addressed: field-level encryption, SSE and selective partial decryption can be architecturally combined in a principled manner into a coherent multi-phase forensic lifecycle, with each component covering specific NIST and LINDDUN obligations, through AES-GCM field protection, tokenised SSE indexes, attribute-bound decryption policies and SHA-256 hash-chained audit logs.

The third objective needed to show that Design Science Research (DSR) is an appropriate way to justify and validate the PPHF framework as a scholarly contribution. Chapter 3 set up the chain of the positivist-deductive-DSR approach and showed the DSRM process model in all the dissertation chapters and also synthetically data use was justified as being both ethically necessary and methodologically sound. The author [5] found that explicit DSRM phase mapping and evaluation against external standards (LINDDUN and the NIST Privacy Framework) meet expectations of DSR for rigour and relevance. Chapter 6 then closed the DSR evaluation loop by showing that the PPHF artefact can help to mitigate concrete threats identified by LINDDUN and how it can be mapped to NIST Identify-P, Control-P and Protect-P functions to support healthcare privacy risk management. RQ3 is answered: DSR is not just a methodological convenience but the epistemologically correct research strategy for an investigation, the primary contribution of which is the absence of an existing artefact, and so the artefact it is designed to be useful for is evaluated within the framework of DSR.

The fourth objective involved evaluating the PPHF framework in relation to the threat models of LINDDUN in order to determine how privacy resilient it is. In Chapter 6 a systematic LINDDUN walkthrough was done on the PPHF DFD, categorizing the framework's response to the seven threat categories. Linkability and Identifiability threats are reduced at field level in Phases 2, 3 and 4 via field-level encryption, tokenised SSE and selective decryption, however, there are always some residual risks in the scoping phase where field-level structural identifiers are required.. Cryptographic policy enforcement is the means of structurally reducing the threats of non-compliance. Wuyts et al. (2024) conclude that LINDDUN is an evaluative instrument that is mature and reusable and Vieira da Silva et al. (2025) find that the use of LINDDUN in healthcare threat modelling is methodologically appropriate. RQ4 is answered: PPHF shows systematic mitigation in all seven LINDDUN threat categories, the most comprehensive mitigation in Disclosure and Identifiability and stated residual risks in pattern level Linkability and key management dependencies.

The fifth objective involved synthesising outcomes to evidence-based conclusions and actionable recommendations that related to the forensic capability versus patient privacy trade-off. As shown in Chapter 6, privacy benefits (such as a reduction in secondary exposure, cryptographic support for minimum necessary, and better audit accountability) are real and forensic utility remains for a well-scoped investigation. Chapter 6 of the NIST Privacy Framework mapping table offers specific examples of alignment with Identify-P, Control-P and Protect-P, which then helps to meet HIPAA and GDPR compliance requirements. The answer to RQ5 is that the synthesised outcomes demonstrate that PPHF is useful in practice and is clearly regulation aligned, thus, it is a defensible operational approach to the forensic-privacy conflict, which is missing from the current healthcare organisations.

7.2 Evidence-Based Recommendations

The following recommendations are based on the findings from Chapters 2–6: adopt field-level encryption as a design requirement for all EHR systems and for audit log infrastructure (implemented in compliance with NIST FIPS 197, a standard which is not yet in effect at the time of writing); incorporate PPHF-style scoped evidence profiling into incident response procedures, replacing default full-database imaging with structured minimum-necessary

specifications, documented and legally authorised prior to forensic acquisition; implement the PPHF audit and chain-of-custody layer as a mandatory component of any forensic or security investigation to enable ex post review of evidence access and support accountability obligations under HIPAA § 164.312(b) and GDPR Article 5(2); and map existing privacy programmes to NIST Privacy Framework functions, particularly Control-P and Protect-P, using the PPHF regulatory alignment table as a starting reference to identify gaps in current forensic practice.

For forensic practitioners and investigators, it is recommended to use LINDDUN threat modelling as a preparatory step to be used prior to any healthcare forensic engagement, use the PPHF data flow diagram (DFD) structure to determine which privacy threats exist in the target system and which controls should be put in place before evidence acquisition begins, and take the approach of using attribute-bound decryption workflows as the standard for accessing protected health information (PHI) during an investigation, and only decrypting the PHI during the final stage of the investigation for evidentiary confirmation purposes if there is policy approval for such decryption.

To incorporate minimum-necessary cryptographic controls into future updates to digital forensics standards like ISO/IEC 27037, and to conduct pilot deployments of the PPHF framework or similar cryptographically integrated forensic frameworks in NHS or other public health systems, with the NIST Privacy Framework and LINDDUN as both evaluation criteria, establishing an evidence base for policy-level adoption.

7.3 Limitations of the Research

This dissertation research has four main limitations that should be recognized prior to interpreting and applying the results.

First, the PPHF prototype is only run on synthetic datasets that are relatively small. Although Gonzales et al. (2024) and Mendes (2025) confirm the use of synthetic data as a methodologically defensible foundation for prototype research, the prototype has not been evaluated with the volume of EHRs that might be produced in a distributed system, which is quite large (potentially millions of EHRs). The computational cost of AES-GCM field encryption, the construction of the SSE index and the evaluation of the policy would have to be empirically determined prior to deployment.

Secondly, the CP-ABE access control model is not fully implemented by using a cryptographic library but approximated at the application logic level. This implies that the security of the policy check that is implemented is not as strong as the security of true CP-ABE scheme that can offer collusion resistance and cryptographic binding of attributes to keys. The engineering challenge of full CP-ABE deployment is recognised as a challenge by Oladosu et al. (2024) and Zala et al. (2024), and the prototype's approximation is acceptable for demonstration of DSR, but would need to be replaced by a production grade CP-ABE implementation before it could be used in practice.

Thirdly the evaluation of LINDDUN in Chapter 6 is conceptual-analytical and not adversarial. The prototype was never subjected to any red teaming, penetration test or independent security audit. Datta et al. (2022) point out that threat analysis cannot replace empirical security testing against real-world adversarial conditions, but Wuyts et al. (2024) agree that LINDDUN offers thorough design-level threat analysis.

Fourth, the study is done by a single researcher and not multi-institutional validated. Larsen et al. (2020) recognise that DSR studies by a single researcher may be vulnerable to validity threats due to the lack of design iterations and/or peer review in the construction of artefacts. The use of external evaluative frameworks (LINDDUN and NIST) is structured, but would be good to replicate the prototype independently, as well as the framework design independently, to increase confidence in the findings.

7.4 Future Work

The above limitations directly provide four avenues for future research. First, performance and scalability evaluation: a natural extension of this research is to evaluate the performance and scalability of a more complete PPHF prototype (with production-grade CP-ABE, a formally secure SSE scheme like the one proposed in the paper [65], and full AES-GCM field encryption) using a large-scale synthetic or de-identified EHR dataset, and measure latency, throughput, and key management overhead under realistic forensic workloads.

Second, full CP-ABE integration: integration of CP-ABE with the application layer in the form of a CP-ABE library, as proposed by authors [66],[67], would provide the collusion resistance

and verifiable attribute binding required for production deployments in healthcare. This would also allow for multi-authority ABE [76] to be evaluated, for multi-organisation forensic investigations across hospital networks.

Third, adversarial and red-team evaluation: have the PPHF prototype tested by independent security researchers to check whether there are any vulnerabilities that are not identified by the design level analysis by LINDDUN, especially in key management, token leakage and policy misconfiguration scenarios. The paper [78] has determined that adversarial evaluation is the missing piece to threat modelling in healthcare security research.

Fourth, deployment and user study: a PPHF compliant forensic toolset deployed in a simulated or sandboxed healthcare environment (with NHS Digital or similar organisations) would provide empirical evidence of the usability and investigator adoption and real-life privacy outcomes. The feasibility of decentralised forensic readiness frameworks for medical networks has been shown by the author [79], and a PPHF pilot would provide evidence of their operational feasibility, which is needed to adopt them as policy.

8 Personal Reflection

This dissertation has been as much a process of learning to manage my time and study as it has been an academic pursuit into privacy preserving digital forensics. Initially, I underestimated the cognitive load of maintaining a 15,000-word research project along with other academic and personal pursuits, and the work was reactive instead of planned, and in bits and pieces. As time went on, I realised that time management was a problem that I could think of as a design problem: I started to structure my own work in terms of scoped, manageable chunks with well-defined inputs and outputs, just like I do for the PPHF. This chapter is a reflection on the development of my planning, prioritisation and learning throughout the project and what routines were successful, what changes had to be made due to errors, and what learning strategies I plan to take to future projects.

I used to have an attitude of “I’m going to do this much reading per week” and then cramming when a deadline was coming up, so it was pretty relaxed, and I’d get a lot of reading done and then, when it was time to write, I’d cram it all in. This method was similar to the issues of traditional forensic methods – sources being over-collected with no clear purpose and the attempt to impose order on collected sources occurring at the end of the process. I realized that this was leading to a lack of depth and engagement with the challenging cryptographic and legal content, so I changed to a more structured approach with short, focused sessions with clear objectives, like “map LINDDUN threats to PPHF Phase 2” or “add NIST Privacy Framework citations to Chapter 4”. This adjustment helped me make my progress more visible and helped me make time management something I did regularly: I blocked out two-hour chunks of time and made it clear to others that I wasn’t going to be available for non-essential things during that time.

One other learning point was related to the quality of study in these planned blocks. In the beginning of the project, I was passive in reading articles and highlighting, but I didn’t take the time to make sure that I was consistently converting my highlights into structured notes or argument maps. This was a problem when writing Chapter 2 as I was forced to remember the logic of each paper, and do so under time pressure. To this, I developed a more active reading approach: summarised the main point of each key source in one sentence, identified how it supported the PPHF design/evaluation and indicated specific paragraphs that might be used to

make an argument regarding LINDDUN, NIST and/or cryptographic mechanisms. This practice aligned with the guidance that study skills are more about creating meaningful structures with information than just gathering information. Consequently, later chapters (especially Chapters 4, 5 and 6) were not really about 'finding something to say,' but more about putting together what was already understood into sections.

Even with these advances, I found myself at times being a procrastinator and a super-committer, particularly when I had to deal with the dissertation and other assignments and personal responsibilities. I had times when I avoided doing cognitively demanding activities like designing the prototype for SSE or mapping NIST functions to PPHF components because they were cognitively demanding, and I went for easier, but less critical, activities like formatting or tidying up references. In retrospect, I found that sometimes my plans were too idealistic about what it would take to get into doing more complicated things in the future, and that I should be taking smaller steps (e.g., “write one toy Python function and a short paragraph explaining it” instead of “implement SSE”). This micro-task method helped me to overcome the habit of postponing important tasks and also followed the advice of effective time management, which is to be realistic about yourself and make a conscious effort to schedule.

The last, and most important lesson is about planning and flexibility. At the start of the project, I regarded my plan as a blueprint and didn't like it when unpredictable factors (more feedback, other deadlines, or my own problems) arose and threatened my plan. Eventually, I realized that planning was more of an iterative process, with weekly review of what was actually completed, re-planning the remaining work, and recognizing that some activities would have to be eliminated and/or simplified to ensure the quality of the dissertation as a whole. This change reflects the advice that good time management involves structure and flexibility – contingencies and direction. As I finished up Chapters 6 and 7, I was no longer following a fixed schedule, but rather a living plan, and I found this to greatly decrease stress and focus more on what was conceptually most important.

In retrospect this dissertation has thus been a learning experience on two fronts. Academically I gained a lot of experience in incorporating cryptographic privacy controls in digital forensic lifecycles and how to test them using NIST and LINDDUN. In a personal sense, I discovered

how to plan more deliberately, how to study more actively and how to plan more continuously and reflectively as a practice and not just an administrative task. For future projects I plan to use a more explicit, scoped, work package approach, to make reading into structured argument notes from the beginning, and to review and tweak my scheduling often, rather than after crises have been identified. If I should choose to continue my research and/or professional endeavors that require ongoing self-directed research within the framework of limitations comparable to those of this dissertation, these skills will be crucial.

9 References

1. Adedigba, A.P., Adewumi, A. and Aibinu, A.M. (2022) 'Securing electronic health records against insider threats: A supervised machine learning approach', *Smart Health*, 26, p. 100354. doi: 10.1016/j.smhl.2022.100354.
2. Al-Dhaqm, A. *et al.* (2022) 'An extended digital forensic readiness and maturity model', *Forensic Science International: Digital Investigation*, 41, p. 301379. doi: 10.1016/j.fsidi.2022.301379.
3. Al-Dhaqm, A. *et al.* (2024) 'Digital forensics readiness framework (DFRF) to secure database systems', *Engineering, Technology and Applied Science Research*, 14(2), pp. 13732–13740. doi: 10.48084/etasr.7116.
4. Al-Hawawreh, M. *et al.* (2024) 'Privacy preservation of electronic health records in the modern era: A systematic survey', *ACM Computing Surveys*, 56(10), pp. 1–36. doi: 10.1145/3653297.
5. Al-mugern, R. *et al.* (2024) 'A cloud forensics framework to identify, gather, and analyze cloud computing incidents', *Engineering, Technology and Applied Science Research*, 14(3), pp. 14483–14491. doi: 10.48084/etasr.7185.
6. Alexei, A. (2022) 'Implementing design science research method to develop a cyber security framework for HEIs in Moldova', in *Proceedings of the 11th International Conference on Electronics, Communications and Computing (IC|ECCO-2021)*. Chişinău: Technical University of Moldova, pp. 1–8.
7. Awolaye, J. *et al.* (2025) 'Forensic analysis frameworks for encrypted cloud storage investigations', *International Journal of Computer Applications*, 187(17), pp. 8–19. doi: 10.5120/ijca2025925241.
8. Carmi, L., Zohar, M. and Riva, G.M. (2023) 'The European General Data Protection Regulation (GDPR) in mHealth: Theoretical and practical aspects for practitioners use', *Medicine, Science and the Law*, 63(1), pp. 22–31. doi: 10.1177/00258024221118411.
9. Carboni, A. *et al.* (2023) 'Privacy by design in systems for assisted living, personalised care and wellbeing: A stakeholder analysis', *Frontiers in Digital Health*, 4, p. 934609. doi: 10.3389/fdgth.2022.934609.

10. Chaddad, A., Wu, Y. and Desrosiers, C. (2024) 'Federated learning for healthcare applications', *IEEE Internet of Things Journal*, 11(5), pp. 7339–7358. doi: 10.1109/JIOT.2023.3337017.
11. Chaure, S. and Mane, V. (2023) 'Digital forensic framework for protecting data privacy during investigation', *EAI Endorsed Transactions on Scalable Information Systems*, 11(2). doi: 10.4108/eetsis.4002.
12. Cormack, A. (2021) 'Thinking with GDPR: A guide to better system design', *Information Services and Use*, 40(4), pp. 321–337. doi: 10.3233/ISU-210107.
13. Deebak, B.D. *et al.* (2022) 'AI-enabled privacy-preservation phrase with multi-keyword ranked searching for sustainable edge-cloud networks in the era of industrial IoT', *Ad Hoc Networks*, 125, p. 102740. doi: 10.1016/j.adhoc.2021.102740.
14. DeSmet, C. and Cook, D.J. (2021) 'Recent developments in privacy-preserving mining of clinical data', *ACM IMS Transactions on Data Science*, 2(4), p. 28. doi: 10.1145/3447774.
15. Dhanalakshmi, G. and George, G.V.S. (2023) 'Secure and privacy-preserving storage of e-healthcare data in the cloud: Advanced data integrity measures and privacy assurance', *International Journal of Engineering Trends and Technology*, 71(10), pp. 238–253. doi: 10.14445/22315381/IJETT-V71I10P222.
16. Etemadi, N. *et al.* (2024) 'Evaluation criteria and methods in design science research: A systematic literature review', *Annals of Computer Science and Information Systems*, 39. doi: 10.15439/2024F7208.
17. Gonzales, A. *et al.* (2024) 'Synthetic data generation methods in healthcare: A review on open-source tools and methods', *Computational and Structural Biotechnology Journal*, 23, pp. 2182–2194. doi: 10.1016/j.csbj.2024.05.007.
18. He, Z. (2023) 'From privacy-enhancing to health data utilisation: The traces of anonymisation and pseudonymisation in EU data protection law', *Digital SocietyDigital Society*, 2(2), p. 17. doi: 10.1007/s44206-023-00043-5.
19. HHS Office for the Assistant Secretary for Preparedness and Response (2023) *2022 healthcare cybersecurity year in review and a 2023 look ahead*. Washington, DC: U.S. Department of Health and Human Services. Available at: <https://www.hhs.gov/sites/default/files/2022-retrospective-and-2023-look-ahead.pdf> (Accessed: 27 May 2026).

20. HIPAA Journal (2024) *Healthcare data breach statistics*. Available at: <https://www.hipaajournal.com/healthcare-data-breach-statistics/> (Accessed: 27 May 2026).
21. Hu, X. and Du, Y. (2024) 'Securing medical data: The integration of Advanced Encryption Standard and blockchain', *Journal of Intelligent Applications*, 2024(1). doi: 10.53964/jia.2024001.
22. ICO (2021) *Data minimisation: Guide to the General Data Protection Regulation (GDPR)*. Wilmslow: Information Commissioner's Office. Available at: <https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/principles/data-minimisation/> (Accessed: 27 May 2026).
23. imec-DistriNet, KU Leuven (2023) *LINDDUN: A framework for privacy threat modelling*. Available at: <https://linddun.org> (Accessed: 27 May 2026).
24. Jaya Vijaya, B. *et al.* (2024) 'Chain of custody practices in modern digital forensics', *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, 10(6). doi: 10.32628/CSEIT2410612443.
25. Kachurka, V. *et al.* (2023) 'Using design science research method to develop a cybersecurity framework', *Annals of Computer Science and Information Systems*, 30, pp. 1–8. doi: 10.15439/2022F7208.
26. Larsen, K.R. *et al.* (2020) 'Validity in design science research', in Hofmann, S., Müller, O. and Rossi, M. (eds.) *Designing for Digital Transformation*. Lecture Notes in Computer Science, vol. 12388. Cham: Springer, pp. 272–282. doi: 10.1007/978-3-030-57993-7_20.
27. Malik, A.W. *et al.* (2024) 'Cloud digital forensics: Beyond tools, techniques, and challenges', *Sensors*, 24(2), p. 433. doi: 10.3390/s24020433.
28. Mendes, J.M. (2025) 'Synthetic data generation: A privacy-preserving approach to accelerate rare disease research', *Frontiers in Digital Health*, 7, p. 1563991. doi: 10.3389/fdgth.2025.1563991.
29. Mpungu, C., George, C. and Mapp, G. (2024) 'Enhancing digital forensics readiness in big data wireless medical networks: A secure decentralised framework', *arXiv preprint*, arXiv:2408.00568. doi: 10.48550/arXiv.2408.00568.

30. Munjal, K. and Bhatia, R. (2022) 'A systematic review of homomorphic encryption and its contributions in healthcare industry', *Complex and Intelligent Systems*, 9, pp. 3759–3786. doi: 10.1007/s40747-022-00756-z.
31. Murtaza, H. *et al.* (2023) 'Harnessing the power of synthetic data in healthcare: Innovation, enhancement and limitations', *npj Digital Medicine*, 6(1), p. 188. doi: 10.1038/s41746-023-00927-3.
32. Narasimhan, P. and Kala, N. (2024) 'Ensuring the integrity of digital evidence: The role of the chain of custody in digital forensics', *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, 10(6). doi: 10.32628/CSEIT2410612443.
33. Naresh, V.S. and Thamarai, M. (2023) 'Privacy-preserving data mining and machine learning in healthcare: Applications, challenges, and solutions', *WIREs Data Mining and Knowledge Discovery*, 13(3), e1490. doi: 10.1002/widm.1490.
34. NIST (2020) *Privacy framework: A tool for improving privacy through enterprise risk management, Version 1.0*. Gaithersburg, MD: National Institute of Standards and Technology. Available at: <https://www.nist.gov/privacy-framework> (Accessed: 27 May 2026).
35. NIST (2021) *LINDDUN privacy threat modeling framework* [informative resource]. Gaithersburg, MD: NIST. Available at: <https://www.nist.gov/privacy-framework/linddun-privacy-threat-modeling-framework> (Accessed: 27 May 2026).
36. NIST (2022) *Hash functions: NIST policy on hash functions*. Gaithersburg, MD: NIST Computer Security Resource Center. Available at: <https://csrc.nist.gov/projects/hash-functions/nist-policy-on-hash-functions> (Accessed: 27 May 2026).
37. NIST (2024) *Cybersecurity framework 2.0*. Gaithersburg, MD: National Institute of Standards and Technology. Available at: <https://www.nist.gov/cyberframework> (Accessed: 27 May 2026).
38. NIST Computer Security Resource Center (2015) *FIPS PUB 180-4: Secure Hash Standard (SHS)*. Gaithersburg, MD: NIST. doi: 10.6028/NIST.FIPS.180-4.

39. NIST Computer Security Resource Center (2001) *FIPS PUB 197: Advanced Encryption Standard (AES)*. Gaithersburg, MD: NIST. Available at: <https://csrc.nist.gov/publications/detail/fips/197/final> (Accessed: 27 May 2026).
40. Obeidat, I. *et al.* (2024) 'Access control solutions in electronic health record systems: A systematic review', *Computers in Biology and Medicine**Computers in Biology and Medicine*, 174, p. 108427. doi: 10.1016/j.compbiomed.2024.108427.
41. Odusami, M. *et al.* (2024) 'Enhancing smart healthcare networks: Integrating attribute-based encryption with blockchain', *Heliyon**Heliyon*, 11(1), p. e39462. doi: 10.1016/j.heliyon.2024.e39462.
42. Oladosu, J. *et al.* (2024) 'Comparison of attribute-based encryption schemes in securing electronic health records', *Scientific Reports**Scientific Reports*, 14, p. 7029. doi: 10.1038/s41598-024-57692-w.
43. Osei, B. *et al.* (2023) 'A detection and investigation model for the capture and analysis of network crimes', *Engineering, Technology and Applied Science Research**Engineering, Technology and Applied Science Research*, 13(5), pp. 11657–11663. doi: 10.48084/etasr.6316.
44. Rahimian, M. *et al.* (2023) 'Cloud evidence tracing system: An integrated forensics investigation system for large-scale public cloud platform', *Forensic Science International: Digital Investigation*, 44, p. 301482. doi: 10.1016/j.fsidi.2022.301482.
45. Randles, M. *et al.**et al.* (2020) 'The positivism paradigm of research', *Academic Medicine**Academic Medicine*, 95(5), pp. 690–694. doi: 10.1097/ACM.0000000000003093.
46. Rauthan, J.S. (2025) 'Homomorphic encryption in healthcare industry applications for protecting data privacy', *arXiv preprint**arXiv preprint*, arXiv:2501.04058. doi: 10.48550/arXiv.2501.04058.
47. Roumpies, F. and Kakarountas, A. (2023) 'A review of homomorphic encryption and its contribution to the sector of health services', in *Proceedings of the 27th Pan-Hellenic Conference on Progress in Computing and Informatics (PCI 2023)*. New York: ACM, pp. 1–6. doi: 10.1145/3635059.3635096.
48. Sciancalepore, S. *et al.* (2024) 'Digital evidence chain of custody: Navigating new realities of digital investigations', in *Proceedings of the 2024 IEEE International*

Conference on Trust, Privacy and Security in Intelligent Systems, and Applications (TPS-ISA 2024). IEEE, pp. 11–20.

49. SecureWorld (2023) *Privacy by design: Integrating privacy into healthcare systems and processes*. Available at: <https://www.secureworld.io/industry-news/privacy-by-design-healthcare-systems> (Accessed: 27 May 2026).
50. Setiawan, I. *et al.* (2024) 'GDPR compliance of hospital management systems in the UAE', *Journal of Data Science and Intelligent Systems*, 2(2), pp. 71–83. doi: 10.47852/bonviewJDSIS42022747.
51. Singh, B.M. (2023) 'A novel secure authentication protocol for e-health records in cloud with a new key generation method and minimized key exchange', *Journal of King Saud University – Computer and Information Sciences*, 35(9), p. 101629. doi: 10.1016/j.jksuci.2023.101629.
52. Soni, P. *et al.* (2024) 'Blockchain-enabled EHR access auditing: Enhancing healthcare data security', *Heliyon*, 10(16), p. e34407. doi: 10.1016/j.heliyon.2024.e34407.
53. Tabassum, M. *et al.* (2024) 'Anomaly-based threat detection in smart health using machine learning', *BMC Medical Informatics and Decision Making*, 24(1), p. 347. doi: 10.1186/s12911-024-02760-4.
54. Tian, H. *et al.* (2020) 'Public auditing of log integrity for shared cloud storage systems via blockchain', *Wireless Networks*, 27, pp. 2285–2298. doi: 10.1007/s11276-020-02373-5.
55. Unal, D. *et al.* (2021) 'A secure and efficient Internet of Things cloud encryption scheme with forensics investigation compatibility based on identity-based encryption', *Journal of Network and Computer Applications*, 194, p. 103213. doi: 10.1016/j.jnca.2021.103213.
56. UpGuard (2025) *14 biggest healthcare data breaches [Updated 2025]*. Available at: <https://www.upguard.com/blog/biggest-data-breaches-in-healthcare> (Accessed: 27 May 2026).
57. Vieira da Silva, C. *et al.* (2025) 'Threat modeling in healthcare: An analysis of trends, tools and research directions', in *Proceedings of the Brazilian Symposium on Information*

- and Computational Systems Security (ERRC 2025)Proceedings of the Brazilian Symposium on Information and Computational Systems Security (ERRC 2025)*. doi: 10.5753/errc.2025.4639.
58. Wairimu, S. and Fritsch, L. (2022) 'Modelling privacy harms of compromised personal medical data — beyond data breach', in *Proceedings of the 17th International Conference on Availability, Reliability and Security (ARES 2022)*. New York: ACM. doi: 10.1145/3538969.3544462.
 59. Wuyts, K. *et al.* (2024) 'Robust and reusable LINDDUN privacy threat knowledge', *Computers and SecurityComputers and Security*, 142, p. 103903. doi: 10.1016/j.cose.2025.103903.
 60. Xie, H. *et al.* (2024) 'Privacy-preserving medical data collaborative modeling: A differential privacy enhanced federated learning framework', *Journal of Knowledge Learning and Science TechnologyJournal of Knowledge Learning and Science Technology*, 3(4), pp. 340–350. doi: 10.60087/jklst.v3.n4.p340.
 61. Zala, K. *et al.* (2024) 'Designing an attribute-based encryption scheme with an enhanced anonymity model for privacy protection in e-health', *SN Computer ScienceSN Computer Science*, 5(2). doi: 10.1007/s42979-023-02541-2.
 62. Zhao, Y. *et al.* (2023) 'Privacy-preserving healthcare monitoring for IoT devices under edge computing', *Computers and SecurityComputers and Security*, 126, p. 103076. doi: 10.1016/j.cose.2023.103076.
 63. Zuo, Z. *et al.* (2021) 'Data anonymization for pervasive health care: Systematic literature mapping study', *JMIR Medical InformaticsJMIR Medical Informatics*, 9(10), p. e29871. doi: 10.2196/29871.
 64. Abdi, I. (2023) *IoT devices in healthcare: Vulnerabilities, threats and mitigations*. MSc thesis, St. Cloud State University. Available at: https://repository.stcloudstate.edu/msia_etds/139 (Accessed: 27 May 2026).
 65. Al-Hawawreh, M. and Sitnikova, E. (2023) 'Leveraging deep learning models for ransomware detection in the healthcare Internet of Things', in *Proceedings of the 2023 IEEE 21st International Conference on Industrial Informatics (INDIN)*. IEEE, pp. 1–6. doi: 10.1109/INDIN51400.2023.10217968.

66. Alotaibi, B. (2023) 'A blockchain-based framework for privacy-preserving and secure management of health data', *Sensors*, 23(6), p. 3139. doi: 10.3390/s23063139.
67. Alzubi, J.A. *et al.* (2021) 'Intrusion detection system based on hybridizing a modified binary grey wolf optimization and particle swarm optimization', *Expert Systems with ApplicationsExpert Systems with Applications*, 185, p. 115641. doi: 10.1016/j.eswa.2021.115641.
68. Bani Issa, W. *et al.* (2020) 'Privacy, confidentiality, security and patient safety concerns about electronic health records', *International Nursing Review*, 67(2), pp. 218–230. doi: 10.1111/inr.12585.
69. Bhushan, B. *et al.* (2021) 'Unification of blockchain and Internet of Things (BIoT): Requirements, working model, challenges and future directions', *Wireless NetworksWireless Networks*, 27(1), pp. 55–90. doi: 10.1007/s11276-020-02445-6.
70. Celesti, A. *et al.* (2021) 'A federated EHR system based on blockchain and IPFS for privacy preserving and data ownership', *IEEE AccessIEEE Access*, 9, pp. 1–15. doi: 10.1109/ACCESS.2020.3047164.
71. Chen, I.Y. *et al.* (2021) 'Ethical machine learning in healthcare', *Annual Review of Biomedical Data Science*, 4, pp. 123–144. doi: 10.1146/annurev-biodatasci-092820-114757.
72. Dang, L.M. *et al.* (2022) 'A survey on Internet of Things and cloud computing for healthcare', *Electronics*, 11(7), p. 1080. doi: 10.3390/electronics11071080.
73. Datta, P. *et al.* (2022) 'A systematic review of forensic frameworks for cyber incident response', *Forensic Science International: Digital Investigation*, 42, p. 301399. doi: 10.1016/j.fsidi.2022.301399.
74. Draz, U. *et al.* (2022) 'Efficient and privacy-preserving e-health system using attribute-based encryption for IoMT', *IEEE Access*, 10, pp. 114556–114567. doi: 10.1109/ACCESS.2022.3217777.
75. El-Hajj, M. *et al.* (2021) 'A critical review of intrusion detection systems in the Internet of Things: Techniques, deployment strategy, validation strategy, attacks, public datasets and challenges', *Cybersecurity*, 4(1), p. 18. doi: 10.1186/s42400-021-00077-7.
76. European Data Protection Board (EDPB) (2021) *Guidelines 07/2020 on the concepts of controller and processor in the GDPR*. Brussels: EDPB. Available at:

https://edpb.europa.eu/our-work-tools/documents/our-documents/guidelines/guidelines-072020-concepts-controller-and-processor_en (Accessed: 27 May 2026).

77. Falaye, A. *et al.* (2023) 'Privacy-preserving authentication and access control in healthcare IoT environment', *International Journal of Advanced Computer Science and Applications*, 14(3), pp. 1–12. doi: 10.14569/IJACSA.2023.0140301.
78. Gaba, G.S. *et al.* (2022) 'Robust and lightweight key exchange protocol for securing fog-to-things communication in IoMT cyber-physical systems', *IEEE Transactions on Industrial Informatics*, 18(9), pp. 6011–6020. doi: 10.1109/TII.2022.3149833.
79. Grover, J. and Sharma, M. (2022) 'Security issues and challenges in IoT-based healthcare', in *Internet of Things (IoT) Based Smart Systems*. Cham: Springer, pp. 223–238. doi: 10.1007/978-3-030-76653-5_12.
80. Haney, J.M. and Furman, S.M. (2021) 'Toward usable cybersecurity in healthcare: A framework for advancing provider usability of EHR systems', in *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. New York: ACM. doi: 10.1145/3411764.3445648.
81. Hasan, M.K. *et al.* (2021) 'A review on electronic health record systems: Challenges and recent advances', *Digital Health*, 7, p. 20552076211065622. doi: 10.1177/20552076211065622.
82. Ibrahim, A. *et al.* (2021) 'Leveraging machine learning for automatic threat detection in healthcare security systems', in *Proceedings of the IEEE International Conference on Machine Learning and Applications (ICMLA 2021)*. IEEE, pp. 1186–1191. doi: 10.1109/ICMLA52953.2021.00192.
83. Iwendi, C. *et al.* (2021) 'Security of things intrusion detection system for smart healthcare', *Electronics*, 10(12), p. 1375. doi: 10.3390/electronics10121375.
84. Javed, I.T. *et al.* (2021) 'Privacy-preserving blockchain-based healthcare system for IoT devices', *IEEE Transactions on Industrial Informatics*, 18(1), pp. 621–631. doi: 10.1109/TII.2020.3047786.
85. Jiang, F. *et al.* (2021) 'A privacy-preserving attribute-based access control scheme for electronic health records', *IEEE Transactions on Cloud Computing*, 10(4), pp. 2873–2886. doi: 10.1109/TCC.2021.3059685.

86. Khalil, I. *et al.* (2021) 'Privacy-preserving hybrid cloud framework for secure EHR management and monitoring', *IEEE Access*, 9, pp. 93099–93113. doi: 10.1109/ACCESS.2021.3093025.
87. Krishnapriya, S. and Sarath, G. (2021) 'Securing electronic health records using blockchain', *Materials Today: Proceedings*, 37, pp. 2408–2414. doi: 10.1016/j.matpr.2020.08.189.
88. Li, J. *et al.* (2021) 'Searchable symmetric encryption with forward search privacy', *IEEE Transactions on Dependable and Secure Computing*, 18(1), pp. 460–474. doi: 10.1109/TDSC.2020.3025869.

