Journal of
Sensor and Actuator Networks

MDPI

Article

PbDinEHR: A Novel Privacy by Design Developed Framework Using Distributed Data Storage and Sharing for Secure and Scalable Electronic Health Records Management

Farida Habib Semantha [1], Sami Azam [1,*], Bharanidharan Shanmugam [2] and Kheng Cher Yeo [1]

[1] Faculty of Science and Technology, Charles Darwin University, Darwin, NT 0909, Australia
[2] Energy and Resources Institute, Faculty of Science and Technology, Charles Darwin University, Darwin, NT 0909, Australia
* Correspondence: sami.azam@cdu.edu.au

Abstract: Privacy in Electronic Health Records (EHR) has become a significant concern in today's rapidly changing world, particularly for personal and sensitive user data. The sheer volume and sensitive nature of patient records require healthcare providers to exercise an intense quantity of caution during EHR implementation. In recent years, various healthcare providers have been hit by ransomware and distributed denial of service attacks, halting many emergency services during COVID-19. Personal data breaches are becoming more common day by day, and privacy concerns are often raised when sharing data across a network, mainly due to transparency and security issues. To tackle this problem, various researchers have proposed privacy-preserving solutions for EHR. However, most solutions do not extensively use Privacy by Design (PbD) mechanisms, distributed data storage and sharing when designing their frameworks, which is the emphasis of this study. To design a framework for Privacy by Design in Electronic Health Records (PbDinEHR) that can preserve the privacy of patients during data collection, storage, access and sharing, we have analysed the fundamental principles of privacy by design and privacy design strategies, and the compatibility of our proposed healthcare principles with Privacy Impact Assessment (PIA), Australian Privacy Principles (APPs) and General Data Protection Regulation (GDPR). To demonstrate the proposed framework, 'PbDinEHR', we have implemented a Patient Record Management System (PRMS) to create interfaces for patients and healthcare providers. In addition, to provide transparency and security for sharing patients' medical files with various healthcare providers, we have implemented a distributed file system and two permission blockchain networks using the InterPlanetary File System (IPFS) and Ethereum blockchain. This allows us to expand the proposed privacy by design mechanisms in the future to enable healthcare providers, patients, imaging labs and others to share patient-centric data in a transparent manner. The developed framework has been tested and evaluated to ensure user performance, effectiveness, and security. The complete solution is expected to provide progressive resistance in the face of continuous data breaches in the patient information domain.

Keywords: privacy by design; healthcare; electronic health record; blockchain; IPFS; distributed file system; GDPR; cybersecurity

Citation: Semantha, F.H.; Azam, S.; Shanmugam, B.; Yeo, K.C. PbDinEHR: A Novel Privacy by Design Developed Framework Using Distributed Data Storage and Sharing for Secure and Scalable Electronic Health Records Management. *J. Sens. Actuator Netw.* **2023**, *12*, 36. https://doi.org/10.3390/jsan12020036

Academic Editor: Lei Shu

Received: 13 February 2023
Revised: 21 March 2023
Accepted: 27 March 2023
Published: 13 April 2023

1. Introduction

Privacy is becoming increasingly important when it comes to information systems that collect personal and sensitive user data [1]. Developing a regulatory framework to protect an organization's assets from the onslaught of cybercrime is a significant concern for governments everywhere. Most healthcare providers provide customers with online services to access and control their data within the organisation's database. When sensitive information is compromised, it can cause a variety of adverse outcomes, including financial disruption and reputational damage for the victim and the affected organisation. Numerous data privacy threats have been identified [2,3], including unauthorised access, data

theft, data loss, IT incidents, and improper data disposal. Data breach risks are rising every year all around the world, and our research revealed some of the significant statistics on data breaches and the costs associated with them. Over the past 12 months, data breaches had the most significant impact in the cyber security world, where 22 billion records have been publicly exposed [4]. In 2022, personal information was accessed during a data breach at Revolut bank, UK, where more than 50k users were affected around the world, including approximately 20,000 in Europe [5]. The same year, Nelnet, a US-based student loan servicing company, leaked confidential information to over 2.5 million users [6]. The majority of data breaches between January and June 2022 occurred at healthcare service providers in Australia, according to the Office of the Australian Information Commissioner (OAIC) [7]. Optus, an Australian telecommunications company, was attacked by a cyber-attack in September 2022, resulting in 2.1 million customers having one form of personal ID number exposed [8]. On 9 November 2022, one of the largest Australian private health insurance providers, Medibank, stated that around 9.7 million current and former customers and authorised representatives' personal and sensitive healthcare information had been accessed by cybercriminals [9]. Based on this incident, the Australian Federal Police (AFP) declared that this crime could impact millions of Australians and can damage significant Australian businesses [9]. The OAIC's investigation of this data contravention will consider whether Medibank initially implemented practices and procedures to ensure compliance with Australian Privacy Principles (APPs). A failure in violation of Australian privacy law may cost up to AUD 2.2 million for each data breach [10]. Another telecommunications company, Telstra, claims that more than 130,000 personal details of their customers were exposed online after a privacy breach in December 2022 that confronted prominent Australian companies and their reputations [11].

An estimated AUD 40 million is the typical price tag for a data breach involving one million records. Despite the prevalence of data breaches, organisations have not found effective ways to prevent or mitigate their effects. The average data breach cost in the healthcare industry is AUD 408 per record, which is three times the cost in other sectors [12]. Privacy by design gives companies an edge by proactively implementing strong privacy practises into the operations of information systems and business processes, guaranteeing both privacy by default and individual control over their personal data [13]. One of the most challenging problems in software engineering is creating a reliable system, whether for private or business use. Some researchers have proposed methods for discussing and ultimately resolving the problem of data breaches. Some examples are data partitioning techniques, anonymous and pseudonymous systems, the blockchain-based solution, the K-anonymity method, and others. Current data privacy measures still do not do an excellent job of preventing data breaches. More robust privacy measurements need to be considered while proposing a privacy-preserving solution. Privacy principles, standards, impact assessment, and compatibility analysis are valuable resources to be considered while collecting and processing personal data [14–16].

Australian Privacy Principles (APPs) [17] are principles-based laws, and the General Data Protection Regulation (GDPR) [18] is the comprehensive data protection law that establish rules and sets a high standard for data protection. It provides control over personal data and requires organisations to obtain explicit consent from individuals for processing their personal data. GDPR also includes provisions for data breaches and is applied to all organisations that process, store, and manage the personal data of European Union (EU) citizens [19]. Failure to comply with the GDPR can result in significant fines and legal action [20]. The validity of GDPR is widely accepted, as it supports a clear framework for data protection, establishing common standards for all EU member states. There are also data protection laws around the world, with countries adopting similar regulations. In our prior review paper, we conducted a comprehensive systematic literature review to identify and analyse the privacy standards, principles, strategies and limitations of existing data privacy frameworks [21]. Based on our review paper [21], we proposed an initial conceptual framework by analysing the key limitations of privacy by design based

on seven existing frameworks [22]. In this research, we anticipated continuing further comprehensive research on privacy by design and related mechanisms to design a novel privacy by design framework. Based on our proposed framework, we developed the Patient Record Management System (PRMS) prototype by considering the limitations of existing frameworks for secure and scalable electronic health records management.

The main contributions of our work are as follows:

- Our research thoroughly examined twelve existing privacy by design frameworks to extract their key limitations. All identified limitations were integrated to ensure maximum privacy to design and develop our proposed framework 'PbDinEHR' (in Section 3.1.1);
- We integrated three international standards, ISO/IEC 15288, ISO/IEC 29100, and ISO/IEC 27001 and 27002, to design the lifecycle stages, privacy contexts, and security control implementation (in Section 3.1.2);
- We proposed six Healthcare Principles (HPs) compatible with APPs and GDPR to ensure privacy by design for EHR management (in Section 3.2.1);
- In this research, we incorporated privacy design patterns such as dynamic data masking, transparent database encryption, and our proposed HPs to guarantee privacy in each layer of healthcare data collection and processing (in Sections 3.2.2.1 and 3.2.2.2);
- We established compliance between proposed HPs with Privacy Impact Assessment (PIA) (in Section 3.2.3.1) and conducted compatibility analysis with globally verified APPs and GDPR (in Sections 3.2.3.2 and 3.2.3.3);
- We incorporated the Ethereum blockchain and Inter-Planetary File System (IPFS) to create private IPFS and permission blockchain networks to share medical files and ensure secure transactions between healthcare provider organisations (in Section 3.3.4);
- Based on all our proposed privacy by design mechanisms, we developed a functional prototype of PRMS that confirms all possible consequences to establish our proposed framework (in Section 4).

The rest of the paper is structured as follows. The relevant research studies are presented in Section 2. We present the methodology, along with the proposed framework in detail, with three distinct phases (planning, assessment, and implementation), which is by far the most comprehensive portion of the paper, as it details the design and implementation for the overall privacy by design framework, in Section 3. We discuss the evaluation and results in Section 4. The usability and functional testing are presented in Section 5. We present the overall discussion in Section 6. Finally, the conclusion, research limitations, and future research are presented in Section 7.

2. Related Work

Healthcare data breach research initiatives that propose personal and sensitive data privacy are relatively ineffective. To address all these perspectives, a more comprehensive methodology is desirable. We thoroughly investigated the existing frameworks to uncover the necessary components and major flaws to create a trustworthy data privacy research outcome. We selected seven existing frameworks from our prior research [22] that were assessed to identify the fundamental components of privacy by design.

The Victorian public sector suggested a privacy protection framework based on privacy by design principles for public sector organisations. In their research, privacy is embedded into the system design to ensure that personal data are safeguarded from the start [23]. A privacy impact assessment is the essential tool in this framework [24]; hence, privacy design strategies need to be considered in parallel with the privacy principles to proficiently protect from data breaches. Moncrieff et al. [25] proposed a framework that eliminates major roadblocks by discovering healthcare system complications through technology acceptance. Blockchain-based fine-grained access control ensures that only authorised users have access to healthcare data closely related to data ownership [26–29]. The construction of this framework does not state whether any verified privacy standards are incorporated or not to develop this framework. PRIPARE (PReparing Industry to Privacy-by-Design by

Supporting its Application in REsearch) is a framework that supports privacy engineering practices, privacy risk management activities, design strategies, requirement analysis, and compliance [30]. hOCBS is a permission Off-Chain Blockchain System that collaborates with digital wallets, smart contracts and tokens on the application and feature layers to support competent and scalable data control invasion [31]. However, design strategies should be applied to the system development to outline the organisational and technical requirements. A framework to enhance e-Health architecture for privacy and security in the healthcare sector is suggested by Shrestha et al. [32]. This research suggests multi-authority-based access control to protect illegal access to patient personal data. MedBloc is a secure EHR system for sharing and accessing healthcare data that uses a permissioned blockchain network [33,34]. However, there is no indication of integrating the verified privacy mechanisms, which should be considered to establish a competitive privacy-preserving environment in healthcare systems. Perera et al. [35] recommended a privacy by design framework based on a set of guidelines to assess the gaps and capabilities of IOT applications and middleware platforms [36]. Privacy by design fundamental principles and privacy design strategies are a core basis; however, there is no suggestion to comply with the privacy impact assessment to measure the risks and mitigation plan. PISCES (Privacy Incorporated and SeCurity Enhanced Systems) is a privacy by design framework suggested by Foukia et al. [37]. Privacy by design principles are incorporated in the information system operation [37,38]. Indeed, PISCES is mainly grounded on privacy principles, and yet there is no evidence of using any privacy design patterns to ensure an effective privacy-friendly system. The ISO/IEC 29110 basic profile privacy by design in the healthcare sector is a framework based on ISO/IEC (The International Organization for Standardization/International Electrotechnical Commission) 29110 [39]. Fundamental principles of privacy by design and privacy design strategies are unified as standard and functional in the development of this framework [40]. The consequences of implementing this paradigm may not be widespread, as privacy-preserving tools and impact assessments have not measured. We investigated closely related works to address the critical qualities for creating a prolific privacy by design framework. Furthermore, we incorporated five additional innovative frameworks to enhance our research, which are assessed to identify the key limitations, as presented below.

Bari and O'Neill [41] recommended a framework for rethinking patient data privacy in the era of digital health that streamlines the Health Insurance Portability and Accountability Act (HIPAA) by associating it with the European General Data Protection Regulation (GDPR) [42,43] and California Consumer Privacy Act (CCPA) [44,45]. The volume and range of information and data breaches are rapidly expanding; in addition, HIPAA has existed for almost a quarter century [46]. Therefore, the potential to set up clear boundaries and achieve the patient's trust using this HIPAA framework is challenging. Reen, G.S. et al. [47] proposed a framework for a decentralised Patient Centric e-Health Record Management System, using a permission blockchain network and IPFS for file storage, where patients should have control over their health records [48–50]. Healthcare providers should obtain consent when looking for patients' data using Ethereum and smart contracts [51–54]. Tariq et al. proposed blockchain-based fine-grained access control that ensures only authorised users have access to healthcare data closely related to data ownership [26–29]. By applying this framework, emphasising information confidentiality concerns to overcome the challenges is crucial [55]. Cernian et al. [56] proposed PatientDataChain, a framework that aims to provide a healthcare application that allows patients to control their healthcare data. Patients must know what data are collected, where they will be shared and stored, and when the consent expires [57,58]. Still, accumulating patient consent while managing and sharing healthcare data cannot ensure a comprehensive privacy-preserving environment. 'OSHealthRec' is a blockchain-based prototype suggested by Meier et al. [59] that supports fine-grained access controls to safeguard data privacy and security [60]. However, implementing this framework is challenging to accomplish significant consequences, as privacy standards have not been considered. Roehrs et al. [61] suggested a

distributed model named 'OmniPHR' that connects medical data with healthcare providers. Additionally, their research suggested data decentralisation and distributed file systems to share EHR across healthcare organisations [62]. Still, accumulating blockchain and IPFS while managing and sharing healthcare data does not assure a wide-ranging privacy-preserving environment. Thus, privacy standards should be considered when designing a privacy-preserving system.

We identified key limitations for all the above related frameworks. Despite the fact that the related frameworks described above have revealed the critical data privacy sectors, complete solutions for developing a data privacy framework for preserving privacy at all levels of a healthcare system in a distributed environment are still lacking. Based on this relevant work, a summary of where we conducted a comparative analysis is presented in Section 3.1.1.

3. Methodology

We conducted comprehensive research on related studies based on privacy by design to identify and analyses their key limitations using Systematic Literature Review (SLR) [21]. Privacy by design components are identified and assessed by considering and examining many closely related frameworks and mechanisms alongside them (Section 3.1.1). Standards and best practices are examined to identify and determine the controls and contexts to construct the proposed framework (Section 3.1.2). To further simplify and secure data privacy in the Patient Record Management System (PRMS), fundamental Privacy by Design (PbD) principles [63] are hybridized into six Healthcare Principles (HPs) (Section 3.2.1). For patients' privacy to be maintained during data collection, storing, and access, the six healthcare principles (HPs) are applied at each level of data processing. Privacy design strategies such as data-oriented and process-oriented strategies are examined to determine the design patterns that best protect the confidentiality of sensitive data at the system's architecture stages [64]. Dynamic Data Masking (DDM) [65] is a data-oriented design pattern that provides full masking, partial value blurring, email blurring, and random masking based on the data types [66,67]. Likewise, Transparent Database Encryption (TDE) [68] is a data-oriented design pattern supported by access controls and permissions keys to safeguard unauthorised access to personal and sensitive data. Healthcare Principles (HPs) are applied to integrate the process-oriented design patterns, while collecting and processing the EHR. We conducted a compatibility analysis between the proposed HPs with Privacy Impact Assessment (PIA) [24] to identify the risks and design a mitigation plan (Section 3.2.3.1). In order to prove that the suggested healthcare principles (HPs) are completely aligned with the two standards, compliance between the HPs with APPs [17] and GDPR [42] was established (Sections 3.2.2.2 and 3.2.3). However, most of the data (clinical, medicine, treatment, etc.) related to the health sector should be critically secured to ensure data holder (patient) privacy. To provide trust and confidence to the hospitals and clinics in sharing their data, this research also proposed blockchain-based IPFS (Inter-Planetary File System) technology [69] that can be used to share crucial medical data while securing and ensuring patient privacy (Section 3.3.4). All the proposed components were implemented using the secure framework ASP.NET and the SQL Server for managing the database (Section 4). Finally, the developed application was tested by conducting security testing and usability testing to measure the risks (Section 5). In this research, we integrated fundamental mechanisms to develop a privacy-preserving framework that can overcome the identified limitations. The proposed privacy by design framework is presented in detail here.

'PbDinEHR' uses a design-based methodology to translate privacy by design into system requirements, as existing studies have not offered a precise solution. The proposed components are described in processes and subprocesses under three fundamental phases, planning, assessment, and implementation, and the phases are grounded in ISO/IEC 15288 process phases, as shown in Figure 1.

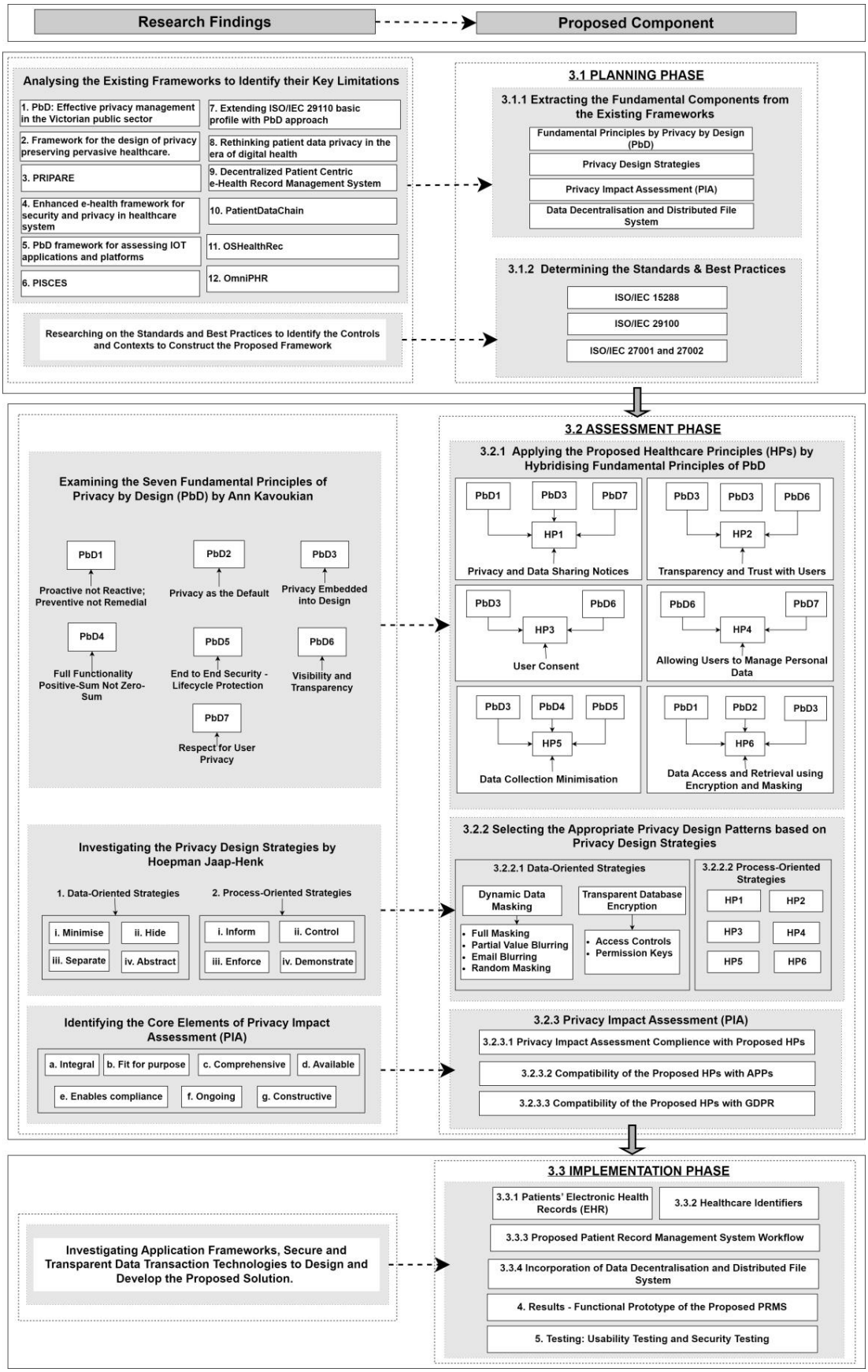


Figure 1. PbDinEHR: A proposed conceptual framework.

3.1. Planning Phase

The planning phase is the first phase that identifies the concerns associated with information privacy so that they can be addressed in the implementation phase. The determination is to characterise the system in terms of privacy perception. In this phase, we identified the critical limitations of privacy by design frameworks. The purpose and application of appropriate standards and best practices are determined to design a privacy defensive EHR.

3.1.1. Extracting the Fundamental Components from the Existing Frameworks

In this paper, we extended our research to analyses of twelve privacy by design frameworks to identify the key characteristics and compare the limitations of individual frameworks. To do so, we highlighted the inadequacies of the selected frameworks and identified four globally validated components. The fundamental components are Ann Cavoukian's seven fundamental principles of Privacy by Design (PbD) [63], Hoepman Jaap-Henk's privacy design strategies [64], Privacy Impact Assessment (PIA) [70] and data decentralisation and distributed file system. Ann Cavoukian's seven fundamental principles of privacy by design are essential components of elementary privacy protection for personal information. Privacy design strategies support the design patterns in the system development life cycle. Data-oriented and process-oriented privacy design strategies deliver the data minimisation techniques for developing a privacy-friendly system. A privacy impact assessment works as a critical component of risk identification and management. PIA addresses the risks associated with the privacy principles and their associated mitigation plan. The focus of this framework was to ensure privacy by design; however, in addition to privacy by design components, blockchain for data decentralisation is included to keep the records of all transactions among the entities to provide transparency [71]. No entity participating in the healthcare environment can access all the transactions in the network. The transactions are only available to the entities participating in the transaction activity. Correspondingly, IPFS for the distributed file system is included for secure file sharing between distributed healthcare organisations [72]. As a result, all these selected components are vital in developing the proposed framework. A comparative analysis between the existing frameworks is presented in Table 1.

Table 1. Comparative analysis of existing frameworks.

Fundamental Components of Privacy by Design (PbD)	**Privacy by Design (PbD) Existing Frameworks**												
	1 [23,24]	**2 [25–29]**	**3 [30,31]**	**4 [32–34]**	**5 [35,36]**	**6 [37,38]**	**7 [39,40]**	**8 [41–46]**	**9 [47–55]**	**10 [56–58]**	**11 [59,60]**	**12 [61,62]**	**PbDinEHR**
Privacy by Design (PbD) Fundamental Principles by Ann Cavoukian [21,63]	●			●	●	●	●			●			●
Privacy Design Strategies by Hoepman Jaap-Henk [64,65]													
• Data-oriented strategies		●	●		●		●		●				●
• Process-oriented strategies		●	●		●		●				●		●
Privacy Impact Assessment (PIA) [24,70,73]	●	●	●			●		●					●
Data Decentralisation and Distributed File Storage [47,71]													
• Blockchain		●	●	●					●		●	●	●

Table 1 presents a comparison of our solution to the existing Privacy by Design (PbD) frameworks. The fundamental components of privacy by design are derived by assessing the relevant works in Section 2. When comparing solutions, we identified that

the existing frameworks do not have at least one or more globally verified components to ensure the privacy contexts, which are limitations for these frameworks. As a result, the feasibilities of the existing frameworks are crucial for achieving the success of personal data privacy management. In Table 1, black dots indicate that the components have been addressed. In contrast, the empty ones indicate that the component is either not addressed or implemented, there is a limitation, or there is still no information provided in the study. We incorporated all of the fundamental components of privacy by design while developing our proposed 'PbDinEHR' for managing patients' EHR in an efficient privacy-friendly environment. The identified components of privacy by design are defined in Assessment Phase. The fundamental components of privacy by design are as follows:

- Seven fundamental principles of privacy by design by Ann Cavoukian [21,63];
- Privacy design strategies by Hoepman Jaap-Henk [64,65];
- Privacy Impact Assessment (PIA) [24,73];
- Data decentralisation and distributed file storage [47,71].

Developing a privacy-preserving framework without integrating all the identified fundamental components of privacy by design can lead to the following limitations. First, when any privacy components are lacking in the design and development of systems and technologies, there is a higher risk of privacy violations; for example, data breaches or unauthorised access to patients' personal and sensitive information [12]. Second, patients and healthcare providers may lose trust if privacy considerations are not prioritised in the design and development, leading to a loss of customers and users and reputational damage [74]. Third, failing to implement privacy by design principles can lead to legal and regulatory non-compliance, as jurisdictions have laws and regulations to implement appropriate privacy protections that can result in legal penalties or consequences. Fourth, it will be more expensive and time-consuming when privacy issues arise after the development of the healthcare system, which can result in higher costs and delays in launching the system or services to market [19,20,43]. Overall, the individual privacy by design component is significant and works collaboratively to ensure privacy while processing EHR; hence, missing any component is considered a limitation when developing a complete privacy-preserving framework. The functionality of the standards and best practices to construct the proposed framework, 'PbDinEHR', are identified here.

3.1.2. Determining the Standards and Best Practices

Standards and best practices were investigated to determine the privacy controls and contexts in the planning phase. Standards were selected based on their purpose for constructing the proposed framework. The processes of the proposed framework and lifecycle stages were established based on ISO/IEC 15288 [75,76]. Based on this standard, three system process phases were created for the proposed framework: planning, assessment and implementation to simplify the design. Privacy contexts and a set of controls for personal data processing were established by following ISO/IEC 29100 [77]. Based on this standard, related privacy measurements were identified and planned to design and develop the proposed framework. As we extended our research, information security management and implementation of the security controls were applied to the proposed framework based on ISO/IEC 27001 and 27002 [78,79]. Based on this standard, information security management controls were established in the proposed framework.

3.2. Assessment Phase

The assessment phase is the second phase that outlines the procedures and development desired to achieve the objectives of the proposed framework. Existing privacy by design frameworks were analysed to extract the fundamental components. Ann Cauvokian's seven fundamental principles of privacy by design are widely assessed in this step. Jeep-Hank Hoepman's privacy design strategies are identified and analysed in order to establish the finest design patterns for personal data minimisation. The Privacy Impact Assessment (PIA) is used to compare the proposed Healthcare Principles (HPs) with the

General Data Protection Regulation (GDPR) [80] and the Australian Privacy Principles (APPs) [17]. These crucial data privacy requirements are recognised by applying these key components of Privacy by Design (PbD) in Electronic Health Records (EHR).

3.2.1. Applying the Proposed Healthcare Principles (HPs) by Hybridising Fundamental Principles of PbD

In the assessment phase, the first step was identifying and analysing Ann Cavoukian's seven fundamental Privacy by Design (PbD) principles [21,63]. To maximise the fortification of a patient's EHR, we created Healthcare Principles (HPs) to expedite the recommended design processes following the seven fundamental principles of PbD. The purposes of privacy by design principles are described as follows.

PbD1 requires privacy by design as a proactive rather than reactive behaviour. This approach endeavours to prevent risks from the initiation that allows privacy-invading events to be predicted and averted before they occur. PbD2 ensures that personal data are automatically and by default protected in any system. PbD3 confirms data privacy incorporation comprehensively and thoroughly using prospective measurements to assess the impact of privacy and reduce the likelihood of data breaches due to misuse, error, or misconfiguration [21,63]. PBD4 provides full functionality by establishing a positive-sum balance between aims and reasonable concerns, rejecting redundant ones such as availability vs. privacy or security. PbD5 ensures that privacy principles are consistently integrated throughout the life-cycle process in EHR systems and that unnecessary data are deleted at the end. PbD6 informs all stakeholders participating in EHR systems that all actions must be visible and transparent to providers and users. PbD7 includes noticeable principles in processes to safeguard the user's privacy by default, such as appropriate notices, alerts, and options while collecting and managing personal data [21,63].

Our prior research [22] initially proposed four Healthcare Principles (HPs) applied in the Patient Record Management System (PRMS) for patient registration. To simplify and modify the design process, six healthcare principles were proposed by hybridising the fundamental principles of privacy by design, which will help to assimilate privacy in each layer of the data processing while designing the proposed system, as shown in Figure 2.

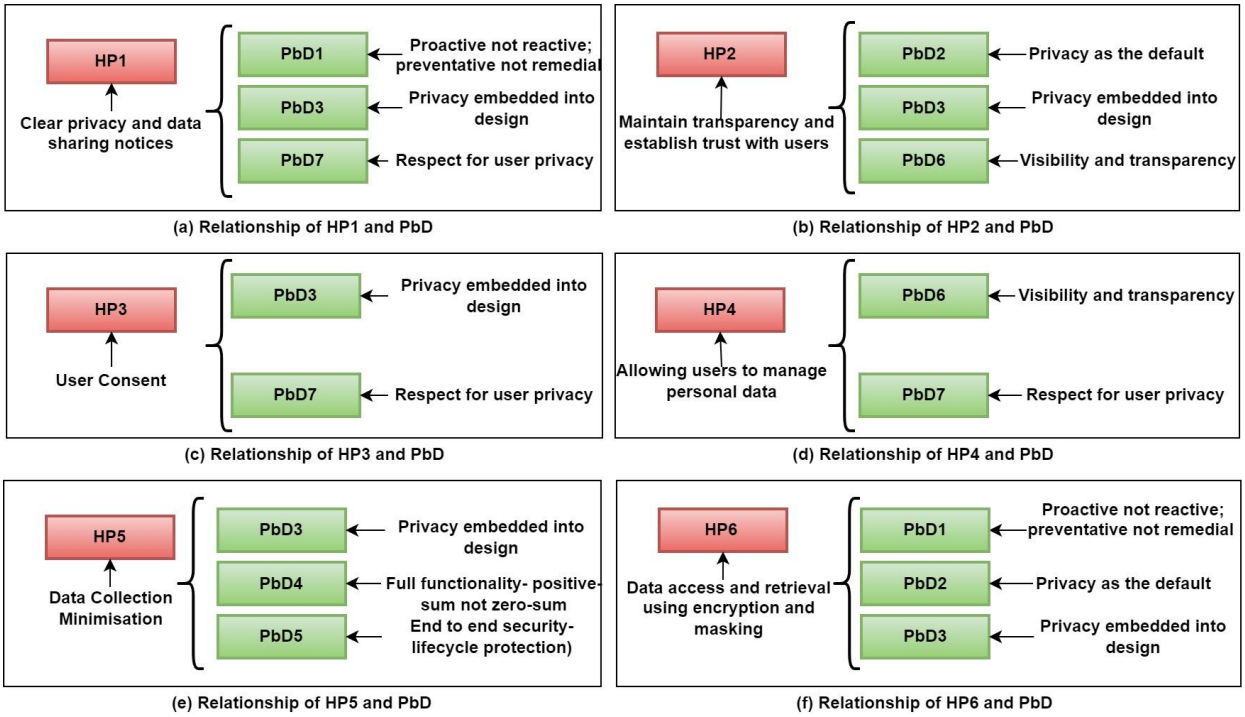


Figure 2. Relationship of proposed HPs and PbD principles.

The proposed Healthcare Principles (HPs) and their relationship to the fundamental principles of Privacy by Design (PbD) are presented in Table 2. These proposed novel principles guarantee maximum privacy in each data processing layer to simplify the design process. The application of these principles will ensure personal data preservation from the beginning. Strong confidentiality and control over personal and sensitive data management will be an additional benefit to healthcare providers.

Table 2. Proposed HPs and their relation to PbD principles.

Healthcare Principles (HPs)	Description of the Proposed HPs	Relationship to PbDs
HP1: Clear Privacy and Data-sharing Notices	HP1 provides customers with explicit privacy and data-sharing notices that explain how their personal information is safeguarded, shared, and deleted. This concept describes the data after the user provides them, whether they will be stored in a database or sent to a third party, and the time limit for data storage. The brief description and data usage policy are established in accordance with the needs of the respective healthcare providers.	PbD1, PbD3 and PbD7 are the foundations of HP1.
HP2: Maintain Transparency and Establish Trust with the Users	HP2 provides notices with an enhanced layer of privacy protection that informs consumers why sensitive data fields are being collected, such as medical reports, laboratory or diagnosis objectives, and so on. When a new user fills out the registration form for the healthcare provider with their personal information, each sensitive data field displays a tooltip or hint for the specific region with necessary privacy notifications. This principle ensures that the healthcare provider maintains transparency and trust with the users.	PbD2, PbD3 and PbD6 are the foundations of HP2.
HP3: User Consent	HP3 ensures that users are notified when a new service accesses their personal information. Before sharing personal information with the new requester, the user must confirm their approval request. Any other significant healthcare notifications will be sent using preferred contact, e.g., mobile, email etc. This concept ensures that the user allows the healthcare provider permission to process the gathered healthcare data.	PbD3 and PbD6 are the foundations of HP4.
HP4: Allowing Users to Perform an Active Role in Managing their Personal Data	HP4 allows users to participate in an active role in personal data management. Users need to read and understand the 'Terms and conditions', which represent the regulations to access, manage, and share personal and sensitive data management and security guidelines.	PbD6 and PbD7 are the foundations of HP3.
HP5: Minimise the Amount of Data Collection	HP5 ensures data minimisation. When the user agrees to the declaration, all the data entered are saved to the cache memory (or temporary memory). The cache memory that holds the data in memory is stored temporarily to minimise the footprint of the actual data. The data in the cache memory will be deleted once the database has been encrypted.	PbD3, PbD4 and PbD5 are the foundations of HP4.
HP6: Data Access and Retrieval by Applying Appropriate Data Masking and Encryption Methods	HP6 ensures the privacy of the acquired data by using Dynamic Data Masking (DDM) and Transparent Database Encryption (TDE). Using these principles, a set of rules and access controls are built. Data collection is secured with appropriate encryption and masking methods based on type, ensuring optimal data gathering. If a healthcare provider requests to see personal data, the data will be retrieved following the healthcare provider's access control policy. If a healthcare provider wishes to alter or update any data, the new data will be acquired using DDM and TDE based on the data categories.	PbD1, PbD2 and PbD3 are the foundation of HP6.

Healthcare principles are vital components of privacy by design; similarly, privacy design strategies ensure that necessary privacy-preserving methods are applied while collecting and managing healthcare data. In the following section, privacy design strategies are assessed to perceive if they should be included in system development during the implementation phase.

3.2.2. Selecting the Appropriate Privacy Design Patterns based on Privacy Design Strategies

Privacy design strategies recommended by Hoepman Jaap-Henk [64] were investigated to design a privacy-protective environment in the PRMS. These strategies recommend design patterns for building a privacy-protected system using appropriate privacy methods. Design patterns support system architects integrating privacy throughout the system development life cycle. There are two categories of privacy design strategies: data-oriented and process-oriented [64].

3.2.2.1. Data-Oriented Strategies

Strategies were assessed to identify suitable patterns to develop the proposed framework. The applications of the data-oriented strategy and the associated design patterns are described as follows.

Minimise is a design strategy that recommends only necessary data be obtained from patients to provide medical services, lowering the danger of data theft, unintended data leakage, and personal data misuse. Individual users can also decide how their data are processed or destroyed when using the system [64]. Hide is a strategy that limits data misuse by properly securing data collection and hiding it from public access while collecting and processing the data legally. Dynamic Data Masking (DDM) is considered the design pattern for these strategies [65,66]. The Separate strategy uses data quality assessment to isolate gathered data and process them secretly. This method safeguards the privacy of EHR, including non-database data such as emails, reports, and system logs. Aggregate ensures that the volume of personal information is controlled and handled with the fewest possible details and a maximum level of combination to make it less sensitive [64]. TDE for access control and permission keys are design patterns for these strategies that allow users to encrypt and control access to an entire database to protect the stored data [68]. In addition, the selected methods, their functionalities, and examples associated with the design patterns are presented as follows.

- *Dynamic Data Masking (DDM)*

Dynamic Data Masking (DDM) is used to hide sensitive data from unauthorised access in a database. Data remain unchanged in the database, but are masked or obscured if retrieved and displayed to unauthorised users. DDM complies with GDPR or HIPAA regulations and protects sensitive information from breaches or accidental exposure. DDM can be applied using four types of mask functions: full masking, partial value blurring, email blurring, and random masking functions [81]. An analysis of DDM and attributes for the functions has been presented in prior research [22]. The patient registration attributes are split to apply the masking functions. DDM limits the sensitive data exposure to users. Unauthorised access to sensitive information is prevented with minimal impact on the application layer. Types of data masking functions [67] and their related examples are defined below:

➢ Attributes for Full Masking

Full masking allows value-making according to the default function's data types. If the data type is a string, values are replaced with XXXX or fewer Xs depending on the field size. If the data type is numeric, values are replaced with Zero values [66,67].

Example SQL Syntax: "[Gender] [varchar] (n) MASKED WITH (FUNCTION = 'default() ') NOT NULL". By using this syntax, the attribute 'Gender' is applied with the default() function and fully masked with 'XXXX' value [67].

➢ Attributes for Partial Value Blurring

This masking method uses a custom string function that includes custom padding in the middle and discloses the first and last letters [66,67].

Example SQL Syntax: "[Healthcare Card No] [nvarchar] (n) MASKED WITH (FUNCTION = 'partial(prefix, [padding], suffix)') NOT NULL". Using this syntax, the attribute

'Healthcare Card no' is applied with a custom padding string in the middle and exposes the first and last letters [67].

➢ Email Blurring

In this method, the email address is masked, but the first letter and constant suffix ".com" is exposed as an email address [66,67].

Example SQL Syntax: "[Email] [nvarchar] (n) MASKED WITH (FUNCTION = 'email()') NOT NULL". By default, this syntax exposes the first letter and constant suffix ".com" in the form of an email address such as aXXX@XXXX.com [67].

➢ Attributes for Random Masking Function

Random masking function is applied to mask any numeric type. The original values are masked with a random value in a specified range [66,67].

Example SQL Syntax: "[Date of Birth] [bigint](6) MASKED WITH (FUNCTION= 'random([start range], [end range])') NOT NULL". The values present in the 'Date of Birth' attribute are masked by applying random masking within a specified range. All selected values are masked in corresponding ranges and syntaxes in random masking [67].

- *Transparent Database Encryption (TDE)*

Transparent Database Encryption (TDE) provides encryption for the entire database. TDE helps to meet regulatory requirements for data protection and reduces the risk of sensitive data being leaked or stolen [82,83]. In our research, TDE is supported by the Microsoft SQL Server to provide encryption for the entire database in a transparent and secure manner. TDE is applied to protect "data at rest" (healthcare data that are stored in the PRMS). 'Master key' and 'Certificates' are created to encrypt the certificates using the master key. User privileges set by the certificates and control mechanisms are used for accessing the database. To encrypt the database, Database Encryption Keys (DEKs) are created for the database users. Therefore, only users with the correct credentials can access the data in the database. Certificates are created to encrypt the DEKs; thus, users with valid credentials can access the specific attributes [68].

➢ Access Controls and Permissions

Database privileges are set up and implemented in SQL Server to determine the users for creating and accessing data stored in the SQL databases [84]. Every SQL Server is securable and associated with permissions that can be granted to the user. Permissions in the Database Engine are managed at the server level assigned to logins and server roles and at the database level assigned to database users and database roles [85]. Server-level roles deliver server-related permissions for creating a new database, and managing logins, backup, shut down and linking to other services, as shown in Figure 3.

	entity_name	subentity_name	permission_name
1	server		CONNECT SQL
2	server		SHUTDOWN
3	server		CREATE ENDPOINT
4	server		CREATE ANY DATABASE
5	server		CREATE AVAILABILITY GROUP
6	server		ALTER ANY LOGIN
7	server		ALTER ANY CREDENTIAL
8	server		ALTER ANY ENDPOINT
9	server		ALTER ANY LINKED SERVER
10	server		ALTER ANY CONNECTION
11	server		ALTER ANY DATABASE
12	server		ALTER RESOURCES
13	server		ALTER SETTINGS

Figure 3. SQL server—permission results.

Database-level roles provide database permissions for accessing the tables. Permissions in the database engine are managed at the server level through logins and server roles and at the database level through database users and database roles mentioned in Figure 4. The model for the SQL Database exposes the same system within each database, but the server-level permissions are not available. A single user can be a member of multiple roles combined with the permissions of different fixed roles to allocate the correct combination based on the requirement [85].

	entity_name	subentity_name	permission_name
1	database		CONNECT
2	database		SELECT
3	database		INSERT
4	database		UPDATE
5	database		DELETE
6	database		VIEW ANY COLUMN ENCRYPTION KEY DEFINITION
7	database		VIEW ANY COLUMN MASTER KEY DEFINITION

Figure 4. SQL database—permission results.

Access controls within the database are essential for the security of data. It is easy to get lost in the jargon of principles, securable, owners, schemas, roles, users, and permissions. Combining the schema-based system (dbo) with various database roles, as shown in Appendix A.1, can provide an easy way to specify permissions for a large collection of securable objects in the database. This allows the creation of several security designs that enable the administrator to restrict access [86].

3.2.2.2. Process-Oriented Strategies

Process-oriented strategies are examined to identify suitable patterns for designing the proposed framework. The application of the strategies and their associated design patterns are described here.

Inform is a process-oriented strategy that ensures that the EHR system informs the users about personal data privacy and the purpose of data collection. Furthermore, if any information needs to be shared with third parties, the patient will be notified and given the opportunity to consent [64]. This technique is implemented using HP1. *Control* ensures that data protection regulations are in place and users have control over their personal information. This strategy is implemented by applying HP3 and HP4 to the EHR system [77,87]. *Enforce* verifies that privacy policy complies with legal requirements during operation, and that procedures are implemented as needed. HP5 and HP6 are design patterns for this strategy, which will be implemented through personal data minimisation, access control, encryption, and masking. *Demonstrate* assures compliance with privacy policies and key public infrastructure. Solid privacy and security measures are beneficial when incorporating vital public infrastructure in healthcare systems. This strategy employs HP2 as a design pattern for auditing, privacy management, and logging activities [77,87].

3.2.3. Privacy Impact Assessment (PIA)

This step protects data by assessing the privacy implications of the proposed healthcare principles. We established a compatibility analysis between our proposed HPs to PIA and constructed a privacy risk assessment and mitigation plan. Compliance between the proposed HPs with the verified standards of APPs and GDPR was established [24,70]. This assessment guarantees that privacy is considered throughout the planning process. When implemented consistently, the PIA prevents and mitigates risks to reduce privacy concerns within the organisation [73].

3.2.3.1. Privacy Impact Assessment Compliance with the Proposed HPs

Table 3 assesses the compliance of the proposed healthcare principles (HPs) with PIA. Since this is a preliminary privacy impact assessment, it is not static and further privacy

implications can be added as needed. The detected threats were analysed, and a risk mitigation plan is produced for each risk in Table 4.

Table 3. Privacy impact assessment compliance with the proposed HPs.

	HP1: Clear Privacy and Data-Sharing Notices	**Y**	**N**	**HPs**
1.1	Does the system shield privacy while sharing and releasing healthcare data?	X		HP1—PbD1, PbD3 and PbD7
1.2	Does the system approve, extract, and release data efficiently?	X		HP1—PbD1, PbD3 and PbD7
1.3	Does the system protect individual data privacy?	X		HP1—PbD1, PbD3 and PbD7
1.4	Does the system send privacy notices while accessing and retrieving personal and sensitive information?	X		HP1—PbD1, PbD3 and PbD7
	HP2: Transparency and trust with the users			
2.1	Are all the collected personal data mandatory to the system?		X	
2.2	Does the system inform users of the reason for collecting and processing their personal and sensitive data?	X		HP2—PbD2, PbD3 and PbD6
2.3	Will the users be reported when collecting their specific personal information?	X		HP2—PbD2, PbD3 and PbD6
	HP3: User consent			
3.1	Does the system send a notification to ask for user consent while managing their personal information?	X		HP3—PbD3 and PbD6
3.2	Do the users confirm the system by approving a notification to use their personal information?	X		HP3—PbD3 and PbD6
3.3	Will the system ask for the user's consent while collecting and sharing healthcare information from one healthcare service to another?	X		HP3—PbD3 and PbD6
	HP4: Allowing users to manage personal data			
4.1	Does the system provide the terms and conditions for storing, sharing, and managing the collected information?	X		HP4—PbD6 and PbD7
4.2	Do the users authorise the terms and conditions as default?		X	HP4—PbD6 and PbD7
4.3	Does the system allow the user to know the timeline of holding their personal information?	X		HP4—PbD6 and PbD7
4.4	Does the system allow users to manage their personal information?	X		HP4—PbD6 and PbD7
4.5	Does the system ask for authorisation from the users if further use or disclosure of personal information is needed outside the original purpose?	X		HP4—PbD6 and PbD7
	HP5: Data collection minimisation			
5.1	Does the system incorporate privacy measurements to ensure the privacy of the collected information?	X		HP5—PbD3, PbD 4 and PbD5
5.2	Does the system ensure the minimisation of collected information before storing it in the database?	X		HP5—PbD3, PbD 4 and PbD5
5.3	Can the user pseudonym themselves when managing their personal information?		X	HP5—PbD3, PbD 4 and PbD5
5.4	Does the system remove unnecessary information once it is no longer required?	X		HP5—PbD3, PbD 4 and PbD5
	HP6: Data access and retrieval using encryption and masking			
6.1	Does the system ask for the user's authorisation to access and retrieve any data?	X		HP6—PbD1, PbD2 and PbD3
6.2	Does the system remove unnecessary information once no longer required?	X		HP6—PbD1, PbD2 and PbD3
6.3	Does the system ensure the minimisation of collected information after applying privacy measurements?	X		HP6—PbD1, PbD2 and PbD3

Table 4. Privacy risk assessment and mitigation plan.

Risk No.	Description of the Identified Risk	Impact	Likelihood	Risk Level	Risk Mitigation Plan	Residual Risk Level
2.1	The system collects personal information that is not compulsory to the healthcare system.	Medium	Low	Medium	The proposed system is for patients with different healthcare service requirements. The system collects personal information that is compulsory for the patients; however, the system has some non-mandatory data fields for the patients that ask the patients to provide information when necessary for treatment purposes. Therefore, some data collection is not compulsory for patients with no prior medical history.	Low
4.2	"Terms and Conditions" are authorised by the user as default.	High	Low	Medium	User acknowledgement is significant when implementing privacy measurements in the healthcare system. The user must read and understand the terms and conditions and must confirm that in the system. Therefore, the user must accept the terms and conditions to verify their authorisation.	Low
5.3	When dealing with data, users will not be able to be anonymous or use a pseudonym	Medium	Low	Medium	As the proposed system is for patient treatments, patients will not be able to mark themselves anonymously. However, healthcare providers will not disclose any information without the patient's consent.	Low

Risk identifier: If any answers reflect No in Table 3, this will be assessed in the 'Privacy Risk Mitigation' in Table 4.

The Office of the Australian Information Commissioner directed potential risks to conduct the impact analysis above. Because the privacy risk assessment generated a low result in the risk mitigation plan, the suggested framework is highly feasible for implementation.

3.2.3.2. Compatibility of the Proposed HPs with APPs

The Australian Privacy Principles (APPs) [17] govern how Australia's personal information is collected and used [88]. Similarly, the General Data Protection Regulation (GDPR) [18] governs how the European Union manages personal information (EU). The compatibility of the proposed healthcare principles and the Australian Privacy Principles (APPs) is presented in Table 5.

3.2.3.3. Compatibility of the Proposed HPs and GDPR

The European Union's General Data Protection Regulation (GDPR) [80] provides guidelines for data protection to support enterprises while collecting, processing, and storing personal data [89,90]. GDPR aims to provide a uniform set of data protection legislation for all EU members, with strict requirements, and allows EU citizens to understand how their data are used and file objections if necessary [19]. Table 6 shows how the suggested principles and GDPR are compatible.

Table 5. Compatibility of the proposed HPs and APPs.

Australian Privacy Principles (APPs)		Purpose of APPs	Compatibility with the Principles of the Proposed Framework
APP1	Open and transparent management of personal information	This principle ensures that personal information is managed openly and transparently with an advanced privacy policy.	HP1, HP2
APP2	Anonymity and pseudonymity	This principle supports data anonymisation and pseudonymisation to protect the user's personal data disclosure.	HP6
APP3	Collection of solicited personal information	This principle provides the management of personal data with an advanced level of privacy measurements.	HP2, HP6
APP4	Dealing with unsolicited personal information	This principle controls unwanted personal information collection.	HP5
APP5	Notification of the collection of personal information	This principle supports notifying the user if the system collects any personal data.	HP3
APP6	Use or disclosure of personal information	The use and disclosure of personal data conditions are outlined in this principle.	HP1, HP4
APP7	Direct marketing	This principle outlines that if any organisation is dealing with a user's personal information, mainly if using and disclosing, they must seek permission from the specific users.	HP3
APP8	Cross-border disclosure of personal information	This principle supports personal data privacy guidelines while disclosing them overseas.	HP4
APP9	Adoption, use or disclosure of government-related identifiers	This principle provides strategies while collecting, using, and disclosing government-related identifiers.	HP4
APP10	Quality of personal information	This principle supports guidelines to maintain the quality of collected personal information. This principle guarantees that the collected data must be correct, up-to-date, and relevant.	HP1, HP4
APP11	Security of personal information	This principle ensures that the user's personal information is secured from loss, misuse and unauthorised access without the user's consent.	HP3, HP6
APP12	Access to personal information	This principle supports appropriate requirements by delivering access to the requests of the users to access the personal information.	HP1
APP13	Correction of personal information	This principle guarantees the correct processes to accurately maintain the user's personal information.	HP1, HP4, HP5

The proposed healthcare principles comply with the Australian benchmark standard Australian Privacy Principles (APPs). Moreover, the General Data Protection Regulation (EU) (GDPR) is an internationally accepted, well-considered, and comprehensive privacy law that recognises personal data's global importance. GDPR is a European Union policy with far-reaching consequences for all enterprises worldwide [91]. We employed APPs and GDPR to assess compliance with our proposed framework because they are standards for measuring when collecting, processing, and storing personal data. Based on the compatibility analysis findings, our proposed principles are fully compatible with the two benchmark requirements, allowing us to ensure the highest level of privacy in patients' healthcare information. The implementation phase is the third phase, and discusses the

proposed Patient Record Management System dataflow in detail. The data decentralisation tool blockchain and distributed file system IPFS are proposed in this phase.

Table 6. Compatibility of the proposed HPs and GDPR.

The General Data Protection Regulation (GDPR)		Purpose of GDPR	Compatibility with Proposed HPs
1	Lawfulness, fairness, and transparency	This principle supports lawfulness, fairness, and transparency in healthcare information. The organisation should have a good reason while processing personal data and ask for consent from the user. The collected data must not be misused and the organisation must be transparent, open, and honest with the data subject and the reason for collecting the user's data.	HP1, HP2, HP3
2	Purpose limitation	This principle sets limitations on using personal data for specific purposes. The data processing boundaries must be established with a notification to the users through a privacy notice. The organisation must limit the data processing to their stated purposes.	HP2, HP3
3	Data minimisation	The GDPR principle of data minimisation suggests avoiding personal data gathering if it is unrelated to the purpose. This principle guarantees that the organisation must collect minor personal data to complete the objectives.	HP5
4	Accuracy	This principle suggests that the organisation should accurately collect and store personal data. They are responsible for setting up regular checks and balances to modify and remove inappropriate and inadequate information accurately. The organisation must have regular basis audits to action removing unnecessary data that are stored.	HP2, HP5
5	Storage limitation	Based on GDPR, the length of time each stored data item is held in a system must be justified. This principle ensures that the data not actively used will be anonymised after a standard time period. This data retention stage helps to meet the storage limitation policy.	HP5
6	Integrity and Confidentiality	GDPR recommends that the organisation maintains the integrity and confidentiality of the personal data collection to keep it secure from internal and external threats. The collected data should be protected with appropriate planning and proactive diligence from unlawful or unauthorised processing and accidental loss or damages.	HP3, HP4, HP6
7	Accountability	The organisation must have proper measures in place as a level of accountability with proof of compliance with the data processing principles. They must have records available at any time that show their compliance with all the rules if managerial authorities ask for this evidence.	HP6

3.3. Implementation Phase

The implementation phase is the third phase, and discusses the proposed Patient Record Management System (PRMS) dataflow in detail. In this phase, the data decentralisation tool blockchain and distributed file system IPFS are proposed. We identified the data and attributes associated with Patients' Electronic Health Records (EHR) and Healthcare Identifiers (HI) that were significant while designing our PRMS workflow. The proposed PRMS prototypes are presented in the results section. Usability and security testing was conducted to measure the outcome and effectiveness of the proposed framework.

3.3.1. Patients' Electronic Health Records (EHR)

Healthcare services depend on the Electronic Health Records (EHR), which assists healthcare professionals and organisations in managing appropriate treatments and services. In Australia, the EHR is mainly stored in a digital system called My Health Record [92]. Personal identification, medical and financial data, and demographic data are all collected as part of a patient's EHR, as presented in Table 7 [93,94]. The healthcare organisation collects patients' personal information to maintain their service registry. During the patient's diagnosis, additional medical information can be added to the EHR by the healthcare providers while the treatment is in operation [95,96].

Table 7. Patients' Electronic Health Records (EHR) [92].

Data	Related Attributes
Personal details	Title, First name, Last name, Date of birth, Gender, Marital status, Healthcare insurance, Occupation, Home address, Street and suburb, State, Phone number, Mobile number, Email
Next of kin details	Name, Phone number, Mobile number, Relationship to you
Emergency contacts	Name, Phone number, Mobile number, Relationship to you
Cultural background information	Cultural background, Country of birth, Is English your first language? Do you require an interpreter? Please specify the language
Allergies and medical information	Allergies and intolerance to medications, Describe your reaction, Regular medication and doses

3.3.2. Healthcare Identifiers

A general scheme for assigning unique identifiers to individuals, healthcare providers, and healthcare provider organisations is implemented based on the Healthcare Identifiers Act 2010 (HI Act) [97]. The Office of the Australian Information Commissioner (OAIC) [98] regulated the privacy aspects of the Healthcare Identifiers Act 2010 (HI Act) [97] and Healthcare Identifiers Regulations 2010 (HI Regulations) [99]. The Healthcare Identifier Service (HI Service) allows healthcare providers to access unique patient healthcare identifiers to match the correct records and maintain accuracy while sharing healthcare information with other healthcare providers. There are three types of Healthcare Identifiers, as follows [98]:

- Individual Healthcare Identifiers (IHI): IHI is for individuals receiving healthcare services, e.g., patients. IHI supports healthcare providers to communicate accurately and identify and access patients' healthcare records [98];
- Healthcare Provider Identifier—Individual (HPI-I): HPI-I is for individual healthcare providers, e.g., Doctors/GPs, specialists, allied health professionals, nurses, dentists, and pharmacists [97,99];
- Healthcare Provider Identifier—Organisation (HPI-O): HPI-O represents organisations providing healthcare services, such as hospitals, clinics, general practices and pathology [97,99].

3.3.3. Proposed Patient Record Management System Workflow

Our proposed healthcare principles (HPs), privacy design patterns, private IPFS and permissioned blockchain network were executed into the proposed framework to ensure privacy while processing patient information. This research mainly selected PRMS to implement the proposed mechanisms. In Figure 5, the workflow represents the entire process from the users', the data owners', and the requesters' points of view.

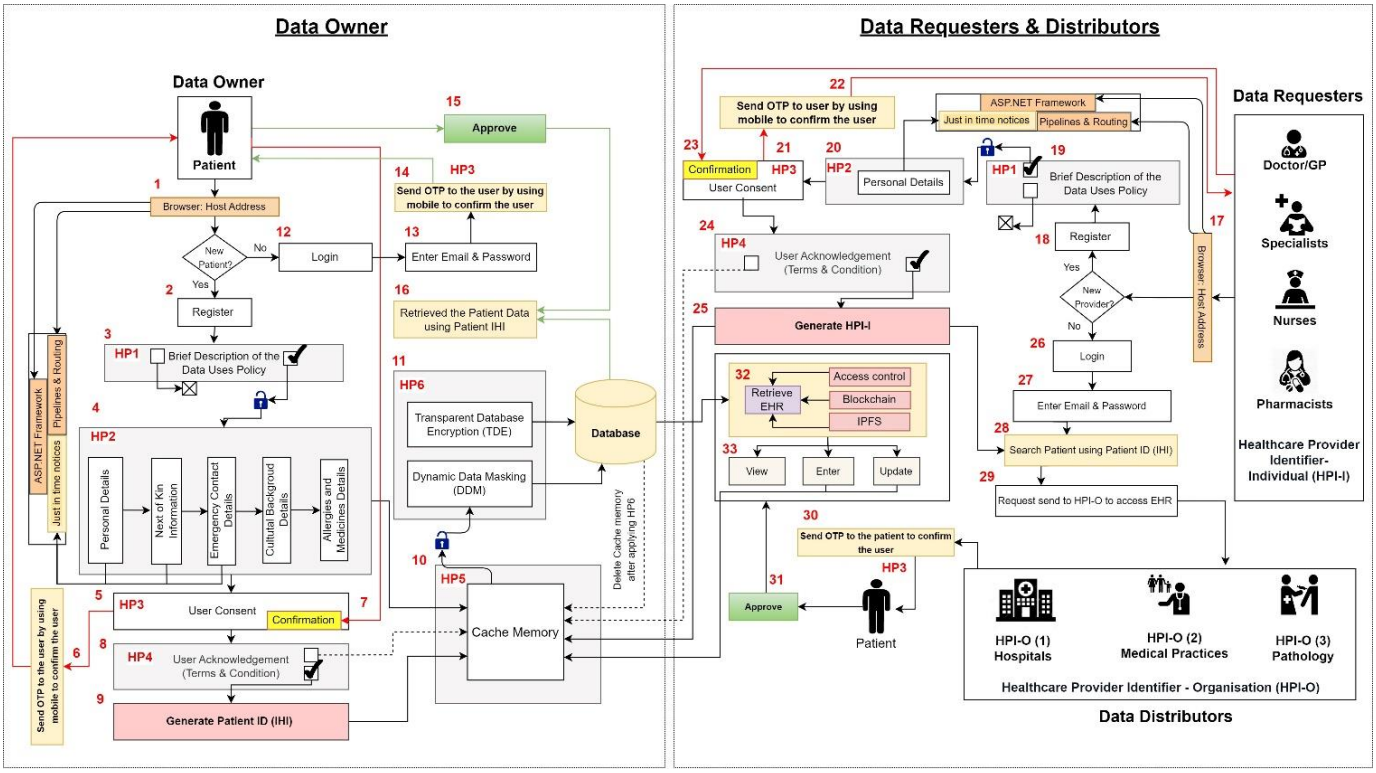


Figure 5. Proposed privacy by design system workflow. Note: Patient (Data Owner) If Not Registered (Step 1–11), Patient (Data Owner) if Registered (Step 12–16), Doctor (Data Requester) if Not Registered (Step 17–25), Doctor (Data Requester) if Registered: (Step 26–33).

The privacy by design system workflow shown in Figure 5 involves three types of users: Data Owners (DOs), Data Requesters (DRs), and Data Distributors (DDs). Dos are the consumers of healthcare services and manage healthcare information. Information stored in PRMS is primarily owned and administered by the DO. In the proposed system workflow, patients are the DO and participate in an active role in managing their personal information by confirming and acknowledging the data uses' policies, terms, and conditions. In contrast, DRs are the individual healthcare providers, such as doctors/GPs, specialists, nurses, dentists, pharmacists, pathologists, etc. In this research, the doctor is selected as the DR who sends requests to the DD to access the patient's healthcare records. DDs are the healthcare provider organisations, such as hospitals, medical practices, pathology, etc. When DRs want to access medical documents such as medical images and test reports, DDs ask for the DOs' approval to retrieve the requested data from the DRs. The workflow of the proposed PRMS integrating the Healthcare Principles (HPs) is described as follows:

3.3.3.1. Patient (Data Owner) If Not Registered

Step 1: A new patient connects to the web browser for the healthcare provider to access the EHR. The healthcare provider's application is designed using the ASP.NET framework.

Step 2: If patients connect with a new healthcare organisation instead of their regular healthcare service providers, they must register with the system first.

Step 3: The patient (DO) starts registering as a new user with the patient interface by accepting that they read the brief description of the data use policy. This step is established based on HP1, which provides clear privacy and data sharing notices to the patients. To start the registration process, the patient needs to read and understand the data uses policy and accept it. The following registration sections will only be available when the patient provides their approval.

Step 4: Once the patient approves, the rest of the section will be activated for them to fill in their details. The patient needs to fill in all mandatory fields marked with '*' in the

registration form. The registration sections are "Personal Details", "Next of Kin Information", "Emergency Contact Details", "Cultural Background Details", and "Allergies and Medicines Details". HP2 is applied to particular data fields that collect sensitive personal data and will display a 'just in time notices' tooltip or hint while collecting the information. The tooltip message is "The collection is necessary for research or statistical activities relevant to public health or public safety, or the management, funding or monitoring of a health service."

Step 5: The DO needs to approve the following section, 'User Consent'. HP3 is applied to this section, while accumulating actual users' consent to manage their personal information and maintaining a healthcare reminder system.

Step 6: A 'one-time password (OTP)' will be sent to the DO for approval. If the patient cannot accept the OTP due to a medical condition, the 'next of kin' registered by the patient can also approve the consent.

Step 7: The DO must approve the OTP request to provide confirmation.

Step 8: The following section is User Acknowledgement. HP4 confirms that an entire 'Terms and Conditions' document is available that allows the patient to manage their personal data in a secure manner. To complete the registration process, the DO needs to accept the 'Terms and Conditions'.

Step 9: After completing all the necessary sections in the registration, a unique patient ID is generated in the system termed "Individual Healthcare Identifier (IHI)".

Step 10: All the entered data in the application interface will be in the browser's cache memory. The cache memory that holds the data is temporary. The HP5 measures are applied to the data available in the cache memory, and the data are unlocked to go to the database.

Step 11: Once the patient data are unlocked in the cache memory, HP6 measures are applied. Dynamic Data Masking (DDM) is applied to each collected data field based on the data types before storing them in the database. Moreover, Transparent Database Encryption (TDE) is used to secure the whole database by creating privileges and certificates for the employees accessing the database. After the application of DDM and TDE, the collected data are stored in the database and the cache memory is removed.

3.3.3.2. Patient (Data Owner) If Registered

Step 12: If the DO (Patient) is already registered with the application, they can log in using their credentials.

Step 13: DO (Patient) logs in using their registered email and password.

Step 14: A "One-time password (OTP)" will be sent to confirm the user.

Step 15: TheDO needs to approve the OTP to confirm their consent.

Step 16: Once the DO approves the OTP, the system will retrieve the patient data using the Patient's IHI.

3.3.3.3. Doctor (Data Requester) If Not Registered

Step 17: The DR connects to the web browser to access the EHR using the ASP.NET framework.

Step 18: If the DR is connecting with the healthcare provider interface for the first time, DR needs to register with the system first.

Step 19: The DR starts registering as a new user by accepting that they read the brief description of the data uses policy. This step is established based on HP1, which provides clear privacy and data sharing notices. DR must approve that they read and understand the privacy policy to start the registration process.

Step 20: DR need to fill out all mandatory fields marked with '*' in the registration section "Personal Details". HP2 is applied to particular data fields that collect sensitive personal data and will display a 'just in time notices' tooltip or hint while collecting the information.

Step 21: The DR needs to approve the following section, 'User Consent'. HP3 is applied to this section, while accumulating actual users' consent to manage their personal information and maintaining a healthcare reminder system.

Step 22: A 'one-time password (OTP)' will be sent to the DR for approval.

Step 23: The DR must approve the OTP request to provide confirmation.

Step 24: HP4 is applied to the next section, 'User Acknowledgement'. 'Terms and Conditions' are available that allow the DR to manage personal data in a secure manner. The DR needs to accept the 'Terms and Conditions' to complete the registration process.

Step 25: After completing all the necessary sections in the registration, a unique ID is generated in the system named "Healthcare Provider Identifier- Individual (HPI-I)" All the entered data in the application interface will be in the browser's cache memory. The HP5 measures are applied to the data available in the cache memory. The cache memory that holds the data is temporary.

After collecting the data in the cache memory, HP6 measures are applied. Dynamic Data Masking (DDM) is applied to each collected data field based on the data types before storing them in the database. Moreover, Transparent Database Encryption (TDE) is used to secure the whole database by creating privileges and certificates for the employees accessing the database. After the application of DDM and TDE, the collected data are stored in the database and the cache memory will be removed.

3.3.3.4. Doctor (Data Requester) If Registered

Step 26: If the DR is already registered with the application, they can log in using their credentials.

Step 27: DR logs in using the registered email and password.

Step 28: After login, DR will be able to search for a patient using the Patient ID 'IHI' to access the EHR.

Step 29: If the patient is not registered with this healthcare provider, the DR will need to send the request to the DD (HPI-O) for accessing the EHR.

Step 30: The DD will send an OTP to the patient associated with the IHI to confirm their authorisation to access their EHR. This OTP will be sent once the new DD needs access for the first time.

Step 31: The patient must approve the OTP request to provide confirmation.

Step 32: The patient's EHR will be retrieved using blockchain, IPFS and access control.

Step 33: The DR can view patients' EHR based on their credentials assigned in the SQL Server. If the DR has the credential to edit or update the EHR, the updated data will be in the cache memory as per HP5. After collecting the user's data in the cache memory, HP6 procedures of DDM and TDE are applied for the updated EHR.

The EHR is distributed between different healthcare provider organisations while requesting and retrieving the patient's healthcare records. EHR distribution uses IPFS for distributed file sharing and blockchain networks for data decentralisation. The proposed privacy by design system is simplified and presented in a sequence diagram in Figure 6.

3.3.4. Incorporation of Data Decentralisation and a Distributed File System

This research aims to design and develop a privacy by design framework to guarantee maximum privacy for patients' personal data. Data decentralisation and the distributed file system are incorporated to share and ensure the secure transaction of medical data between different healthcare provider organisations. In addition, combining these features will provide more scalability to our proposed privacy by design framework. Most healthcare providers rely heavily on cloud-assisted, centralised data centres for storing and distributing patients' medical images and test reports [100]. In recent years, many researchers have developed frameworks using blockchain and distributed file system networks to demonstrate the efficient exchange of medical records between various providers [101]. We have used the Ethereum blockchain and Inter-Planetary File System (IPFS) to create a private IPFS network and two permissioned blockchain networks to share files and record

event logs between various users of the network [72,102]. We have implemented and tested the networks in a Linux (Ubuntu) environment by following the steps provided in the official documentation for Ethereum and IPFS. Integrating this with the policies and the framework discussed above enables the patients to own their data and allows healthcare providers to share medical records transparently with patient consent.

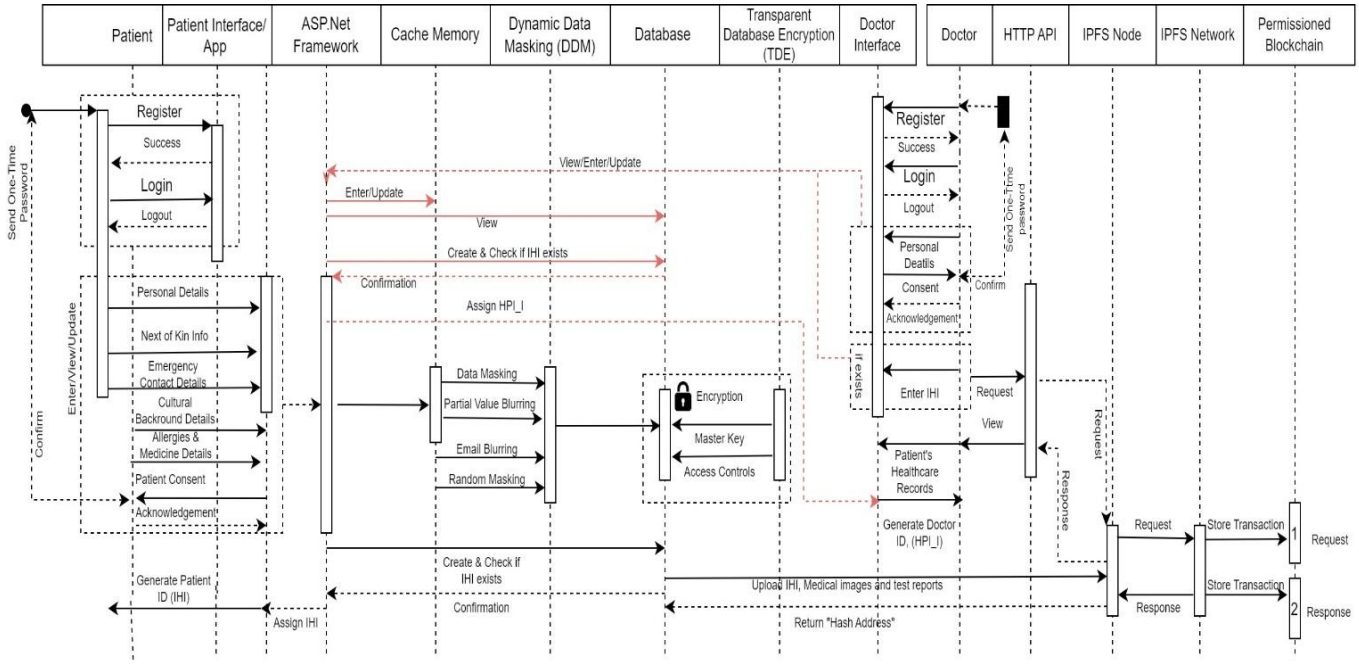


Figure 6. Proposed system sequence diagram.

Figure 7 presents an architectural design of the proposed blockchain and IPFS where HTTP API mainly communicates between the systems. Individual healthcare providers such as a data requester (DR) sends a request to the healthcare provider organisation hospital 1 (HPI-O(1)) to locate the EHR based on the patient's IHI. The healthcare provider will need to request the hash addresses of the patient's medical files. HTTP API gateways are created for all nodes, and the HPI-O nodes are connected to the API gateway of the admin node (node 1). The healthcare provider's admin node API requests medical record hash addresses. This will send a request to all other healthcare provider organisations, such as HPI-O (2) and HPI-O (3), to retrieve the hash addresses for the IHI. The nodes will send the hash addresses (if they exist) to the API gateway of the admin node. After receiving the hash addresses related to the IHI, the admin node will download the medical files.

The following sections will explain the individual aspects of blockchain and distributed file system technologies used in this research to demonstrate the point [102].

3.3.4.1. Consensus

The healthcare provider nodes that are participating in the network should accept the below consensus rules:

- The patients requesting services from the health care providers should have a 'common unique patient ID (IHI)'. This allows for the easy retrieval of the EHR that are stored across various healthcare provider nodes;
- Each healthcare provider will maintain an IPFS node with a unique Node_ID (or HPI-O). These addresses are used to identify the location of the original data;
- The healthcare provider nodes cannot store the IPFS hash addresses on the blockchain network. They are stored locally in their 'information tables'. The hash addresses are shared based on requests between various IPFS nodes (HPI-O nodes);

- The healthcare provider nodes cannot access the blockchain data of other nodes [53];
- Blockchain networks are only used to listen to events in the IPFS network between various nodes and store them on the blockchain ledger, providing improved auditing and transparency [53].

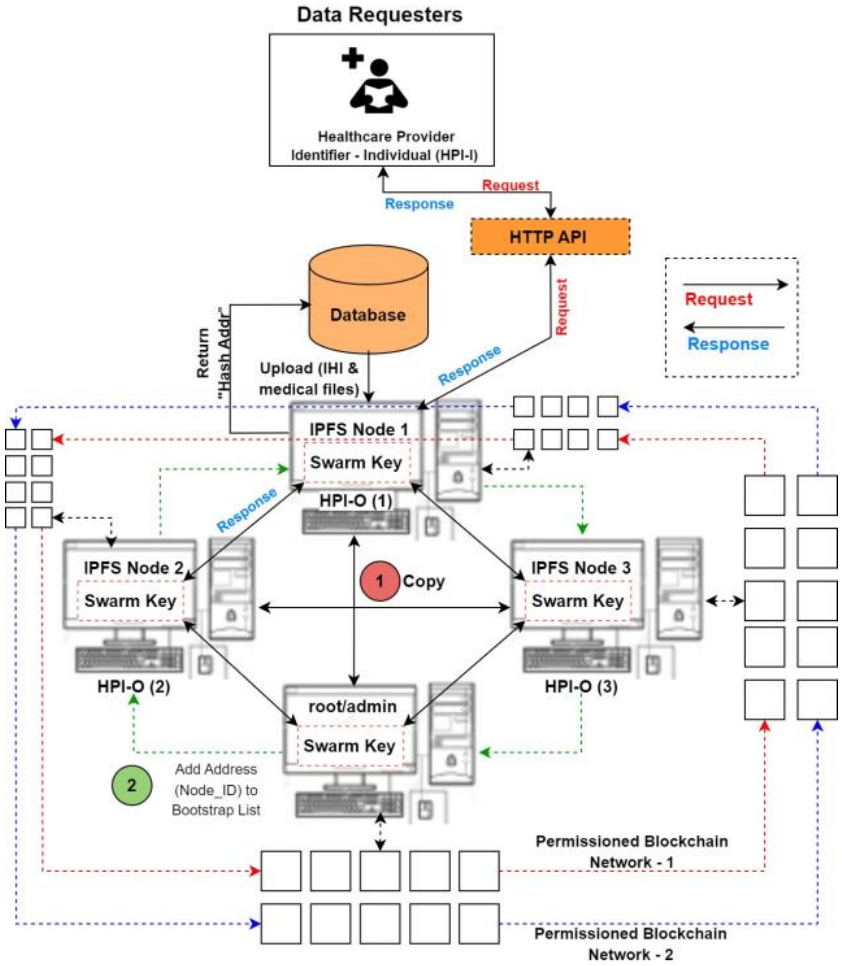


Figure 7. Proposed blockchain and IPFS architectural diagram. Note: DR sends request to HPI-O which is represented as 'Red' and the response is represented as 'Blue'. 1. Coping the Swarm keys marked as 'Red'; 2. Add address (Node_ID) to bootstrap list marked as 'Green'.

3.3.4.2. Private IPFS Network

Private IPFS only allows those nodes that have shared swarm keys. Nodes outside the network cannot communicate with the private network. Before creating the private network between the nodes (healthcare provider organisations), virtual private network (VPN) tunnels should be created in the healthcare settings of various participating nodes. This would allow for secure traffic flow between the IPFS nodes at multiple locations using different internet service providers (ISP). Each node participating in the network should consist of VPN servers that are connected to the VPN server of the admin node [71]. Each healthcare provider organisation (HPI-O) is represented by a node (computer).

These nodes are installed with IPFS-related libraries to initialise IPFS in the node and create a private IPFS network [103]. IPFS nodes in a private network can only communicate with other nodes who share their secret key/swarm key [104]. In the proposed framework, the IPFS node of each healthcare provider are connected to the root/admin IPFS node using shared secret keys, as seen in Figure 8.

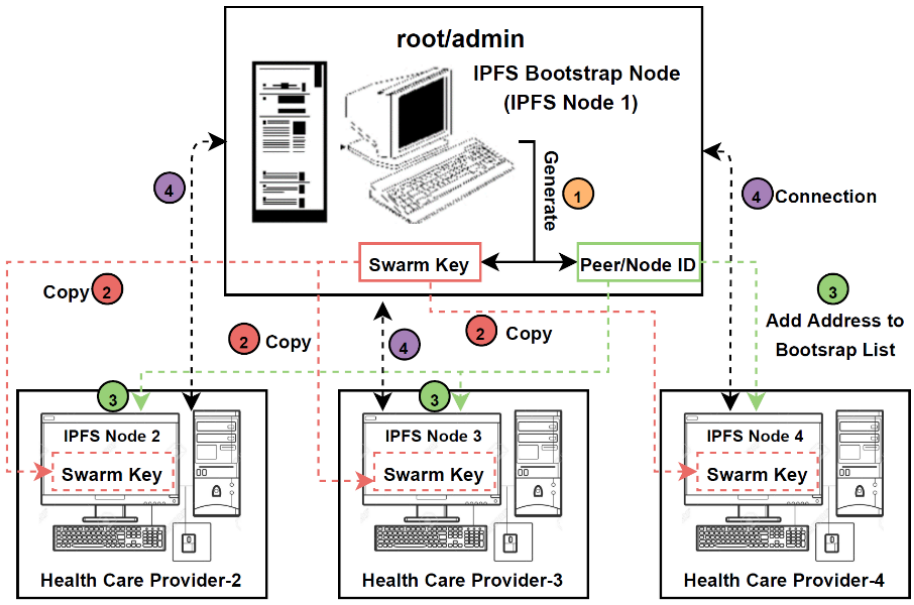


Figure 8. Private IPFS network. Note: 1. Generate swarm key marked as 'Yellow'; 2. Copying the swarm key marked as 'Red'; 3. Add address (Peer/Node ID) to bootstrap list marked as 'Green'; 4. Connection of each node marked as 'Purple'.

The IPFS nodes of the healthcare provider (HPI-O) nodes 2, 3, and 4 are not directly connected. Nodes cannot interact or access medical records without the root node (node 1). This root node can supervise the entire network by validating transactions across healthcare provider IPFS nodes. Any unstable healthcare provider node would prohibit working nodes from accessing its data. To address this issue, we should make duplicates of the original node and use them when a single point of failure happens. As a secondary measure, we can create new nodes to replace outdated or unstable nodes by using the original medical data stored in local SQL databases. The following steps are used to explain the construction and workflow of the proposed private IPFS network:

- Installing IPFS-related libraries on every node;
- Initialising IPFS on every node will create a local IPFS repository in them. The init function generates:
  - A 2048-bit RSA key pair allows the IPFS node to sign the content created on that node cryptographically;
  - A peer ID for the node. Each IPFS node is identified by a unique ID (HPI-O). These HPI-Os are used to create a private network between the nodes.
- Using the 'ipfs-swarm-key-gen' package, a swarm key is generated only in the root node (node 1). SSH or manual transfer is used to copy the admin node's swarm key file into every node participating in the private network that agreed to a consensus. This allows the nodes in the network to communicate with only those nodes that share the same 'secret key';
- Removing default addresses from the IPFS bootstrap list of each node. The IPFS daemons use the addresses added to the bootstrap list to establish a connection with the addresses (nodes). Only the address of the admin node (node 1) is added to the bootstrap list of all the nodes (nodes 2, 3, 4). This results in:
  - Only the admin node having direct access to each IPFS node in the private network;
  - The healthcare providers only accessing each other's medical records through the admin node.

The following instructions are used to add the root node address to the bootstrap list of all the IPFS nodes (Appendix A.2):

root node (IPFS Node 1) address

/ip4/<root node IP>/tcp/<port address>/ipfs/<Peer ID of root node>

Peer ID is the address of an HPI-O node generated after initialising IPFS on the node, and **tcp** is the network protocol. Here, specific port addresses can be used to establish a connection. This address is added to the bootstrap list using the below command on every healthcare provider node (nodes 2, 3, 4).

*# bootstrap add **root node** address onto node 2, 3, 4,*

ipfs bootstrap add/ip4/<*root node IP*>/tcp/<*port address*>/ipfs/<*Peer ID of root node*>

Executing the above command on nodes 2, 3, and 4 connects them to each other and the root node. This can be checked by listing each node's swarmed peers/connections.

The upload process involved the private IPFS network of the proposed framework. The following steps are used to explain in detail the upload process:

- Based on the consensus agreed upon by the healthcare provider organisations participating in the private network, the patients visiting these healthcare settings should have a unique Individual Healthcare Identifier (IHI). This IHI links a patient's medical records at various healthcare settings of the participating nodes;
- If the patient is new among all the participating healthcare provider nodes, then a new IHI is assigned. All the nodes will use these IHIs to identify the patients and their medical records;
- The patient's medical records created by the healthcare providers are stored in their respective local storage devices. These local devices consist of information tables that contain all the related information. The healthcare providers will use their local storage to access their data internally. Here, IPFS is used to provide secure external access;
- Before uploading the medical records to the IPFS node, they are encrypted using the AES-256-bit encryption algorithm. The Advanced Encryption Standard (AES) uses symmetric key encryption, which involves 'one secret key' to encrypt and decrypt the data;
- The medical records are encrypted with the patient's password before uploading to the IPFS, because if someone has the hash address, they can retrieve the file anytime. The medical records are encrypted with the patient's password to tackle this drawback. The medical records can be opened or viewed with consent from the patients or the 'Next to Kin' for new healthcare organisation. Here, we can use 'one-time password' (OTP) features to provide additional patient security;
- After uploading the encrypted medical record onto the IPFS, it is pinned to the respective node. Similarly, the medical records belonging to the health care providers are uploaded and pinned to their individual nodes. It is important to pin the data to the nodes because IPFS nodes treat the uploaded data as a cache, which means there is no guarantee that the uploaded data will stay in the network forever;
- IPFS uses a method called garbage collection to remove data from the nodes if the nodes' disk space is full. If the data are not pinned to the nodes, they might be removed in the future. The IPFS nodes should pin their respective medical records to tackle the problem. They should be pinned if you want the data to be available in the network for the long term;
- Uploading the files to the IPFS nodes generates a hash address, as presented in Table 8. This hash address is generated based on the content in the document;
- The generated hash addresses are returned to the local PC and added to the information table to link the IHIs and medical records with the respective generated hash addresses;
- Finally, the root/admin node (node 1) will pin all the medical records uploaded by various nodes to their nodes. However, if a node hosting some documents goes down, it will be difficult to access the medical records present in that node. To tackle this

drawback, the medical records from all the nodes (nodes 2, 3, and 4) are pinned to the root node.

Table 8. Hash values of the uploaded files.

HPI-O	IHI_ID	HPI-I_ID	Filename	Hash Values of the Uploaded Files
HPI-O (1)	120826	05032503-F759-451B-C733-08DAD850FEE4	Chest X-ray	QmcSrvMHKwfZ6rFwmvyJs2HvHZNZ7Ro2kssfK1JQqXHtjt
HPI-O (2)	120826	05032503-F759-451B-C733-08DAD850FEE4	Blood-Test	QmZsFrynfw5VC956TbeGhrJboizTHnhiB1hjKRRNZBmUuQ
HPI-O (3)	120826	05032503-F759-451B-C733-08DAD850FEE4	MRI-Report	QmTx6se8NVRsnabePDM9bSWPJSSFGuJTYBErebdRktFgBa
HPI-O (4)	122713	F1B848DB-5E53-4BE8-B7A9-08DAE4DC568D	Blood-Test	QmTXbGRE3BaLvHdS8r2kivPgJFhbZsWNfCx1Fv5Hbnthy4
HPI-O (5)	122713	F1B848DB-5E53-4BE8-B7A9-08DAE4DC568D	Blood-Test	QmTSb1WAD66JqANT5Cj94bzs6Y7cPVB4ayEMJhAU6PDa9Z
HPI-O (6)	122713	F1B848DB-5E53-4BE8-B7A9-08DAE4DC568D	Chest X-ray	Qm7MGsyejsowsjSjdhriekjWPJSSFGuJTYBErebdRktgtS5

3.3.4.3. File Request and Response

Each healthcare provider's HTTP API Gateways connect their interface and IPFS node. These connections send patient information requests to the entire IPFS network, receive a list of medical records and addresses (HPI-O/Node) associated with a patient (IHI), and view them through the interface. The following steps are used to explain in detail the process involved in requesting and receiving a list of medical records for a patient and their associated node addresses (HPI-O):

- Patients who request service from a healthcare provider will generate the IHI. This IHI will be used to request the list of medical records that are present across various nodes;
- Running a daemon on an IPFS node exposes an HTTP API, which can be used to control the node. HTTP API gateways are created for all the nodes. The API gateways of all the healthcare provider nodes are connected to the API gateway of the admin node;
- The healthcare provider nodes will use the API connection to request the medical records related to an IHI at various nodes. Figure 9 showcases the addresses that can be connected between the interface and the IPFS nodes;
- The admin node API will use its connections with the APIs of other healthcare provider nodes to request patient information across the network. This information will be sent to the requester node in the form of a list consisting of filenames and data owner addresses (HPI-O);
- Using the HPI-O addresses present in the list, the healthcare providers can use the blockchain network to send request and response transactions for hash addresses of the medical records in the network;
- After receiving the hash addresses related to a patient (IHI), the IPFS node will download the medical records and transmit them to the interface using its API gateways.

Figure 9. HTTP API gateway addresses (Examples).

3.3.4.4. Permissioned Blockchain Networks

The proposed framework consists of two permissioned blockchain networks connected to every node. The process of listening to and storing all the events (upload, request, response, and access) that happen in the IPFS network in a single blockchain network will not be effective. It can affect the auditing process. The admin node (node 1) controls who and what type of data a node can access in the blockchain networks. Blockchain networks also consist of 'smart contracts' that can be used to automate various activities based on predetermined conditions [105]. Using these contracts, the hospitals or clinics can share the data transparently, securely, and efficiently while ensuring patient information privacy. The IPFS and blockchain network entities can manage workflows without any intermediary's involvement or time loss. Transactions among the entities are stored on the blockchain to provide transparency and automation [106,107]. The following steps are used to explain the process involved in our permissioned blockchain network that handles requests. The same steps can also apply for the response blockchain network:

- Every healthcare provider node in the ETH network has a smart contract at their address, which comprises data and functions that can be executed upon receiving a transaction. The state variables (or persistent data) are stored permanently on the blockchain network. Mentioning the data type, as shown in Appendix A.3, allows the contract to keep track of storage on the blockchain;
- Using the emit function, the smart contract can emit events related to request/response transactions. These events are referred to as logs. These logs are written into the blockchain. The structure of the logs for the request network are designed to only store data such as the timestamp, IHI of the patient, and from and to addresses of HPI-O nodes;
- The logs are designed in such a way that their storage in the blockchain should cost less than contract storage. Each healthcare provider node's event logs are stored locally and on the blockchain. This allows for a more outstanding audit of the upload events;
- Similarly, the nodes in permissioned Blockchain Network 2 have their own smart contracts that can store events related to responses from various healthcare provider nodes (HPI-O);
- The log file consists of information such as 'HPI-O of the requester node', 'IHI', 'Hash addresses of the requested files', and 'Timestamp'. The generated log files are stored in the response blockchain network;
- Rather than listening to and storing the events directly on the blockchain through the smart contract, the logs are emitted from the contract and then stored on the blockchain. The admin node will use these logs to perform audits on the data. All the log files can be addressed with the respective contracts of each node. This allows for the easy retrieval of wanted information from the blockchain;
- All the network transactions are cryptographically signed instructions that can be sent between various accounts (HPI-Os) in the Ethereum network. Any statement in the network can initiate a transaction to update the state of the network. These

transactions are broadcasted to the whole network so that a validator can execute the transactions and propagate the changes to the network. A transaction from an account on the ETH network includes the following information.

Ethereum has two account types: externally owned accounts (EOA), and contract accounts that provide the ability to receive, hold, and send transactions and interact with the deployed smart contracts. However, with EOAs, the transactions between these accounts can only be ETH/token transfers and they do not allow accounts to trigger codes that can execute different actions, such as creating a new contract. In our research, to request and receive (transactions) the hash addresses of the files stored at various HPI-O nodes, we connected two blockchain networks/nodes (request and response) to the IPFS network/nodes. This was done so that all the transactions between various accounts could be validated based on an assigned key and stored on a distributed ledger which could be audited using the from and to HPI-O addresses of the requesting/responding node, as shown in Appendix A.4.

As the world is moving towards the new information age, technologies such as IPFS and blockchain play a crucial role. These technologies allow for secure and transparent open-data initiatives for data sharing among various entities. Sharing critical medical records of patients among different healthcare providers can provide better patient care on time and save costs. Moreover, blockchain's immutability clashes with GDPR's "rights to erasure". Blockchain is designed to be immutable, a decentralised and distributed ledger where personal data cannot be deleted or modified, which goes against the GDPR's right to erasure [20]. This research has proposed a framework with a private IPFS and two permissioned blockchain networks. The private IPFS allows the secure storing of medical records of various health care providers (IPFS nodes) on the network. In addition, the two permissioned blockchain networks are used to store events related to file requests and responses. Blockchain technology was carefully considered to balance the immutability with GDPR's right to control personal data. The proposed framework will allow the healthcare provider organisations (HPI-Os) and the patients (IHIs) to own their medical records while sharing them with various healthcare provider nodes that have accepted the consensus rule and participating in the networks. The prototypes of the proposed frameworks are presented in the following results section.

4. Results

This section outlines the results of the developed framework 'PbDinEHR'. The screenshot of the functional prototype is presented in the following section.

4.1. Functional Prototype of the Proposed PRMS

The proposed framework was implemented using technologies and protocols that align with the requirements, such as ASP.NET, Microsoft SQL Server Management Studio 18, Visual Studio 2019, Ethereum blockchain and Inter-Planetary File System (IPFS). ASP.NET was used as it is a user-friendly and secure framework suitable for healthcare applications. ASP.NET provides tools and libraries to simplify the process of building our application. In addition, ASP.NET supports web forms to simplify dynamic web pages with server-side controls, MVC for design patterns, Web API to build HTTP services that help access from web browsers and mobile devices, SignalR for a real-time communication library and Entity framework, which is a robust Object-Relational Mapping (ORM) tool that simplifies the process for working with the databases. ASP.Net is a reliable framework that offers security, scalability, and good performance for building our proposed PRMS application. Microsoft SQL Server Management Studio 18 (SSMS 18) manages and administers SQL Server databases, including creating, modifying, and deleting databases and tables, and managing users and permissions [85]. Visual Studio 2019 is an integrated development environment (IDE) for creating applications using ASP.NET. This tool provides built-in support for creating ASP.NET applications with a range of templates and tools to create and deploy the application. The InterPlanetary File System (IPFS) is a protocol and network

designed to create a permanent and decentralized method for storing and sharing files. Ethereum blockchain is a decentralized and distributed ledger that executes smart contract and records transactions for our proposed PRMS. The InterPlanetary File System (IPFS) is a protocol and network designed to create a private IPFS network and two permissioned blockchain networks for storing and sharing files [101,108]. Examples of the patient and doctor's portal prototype and their associated unmasked and masked data are presented in the below section.

4.1.1. Patient Registration Prototype (Personal Details)

The proposed components are implemented using the application framework ASP.NET and MySQL server. For example, Section A: Personal Details from the patient registration form is presented in Figure 10 with a few of the associated unmasked data in windows authentication, as shown in Figure 11, and masked data in SQL Server authentication, as shown in Figure 12. The creation of the Patient ID (IHI_ID) once the patient is registered is presented in Figure 13.

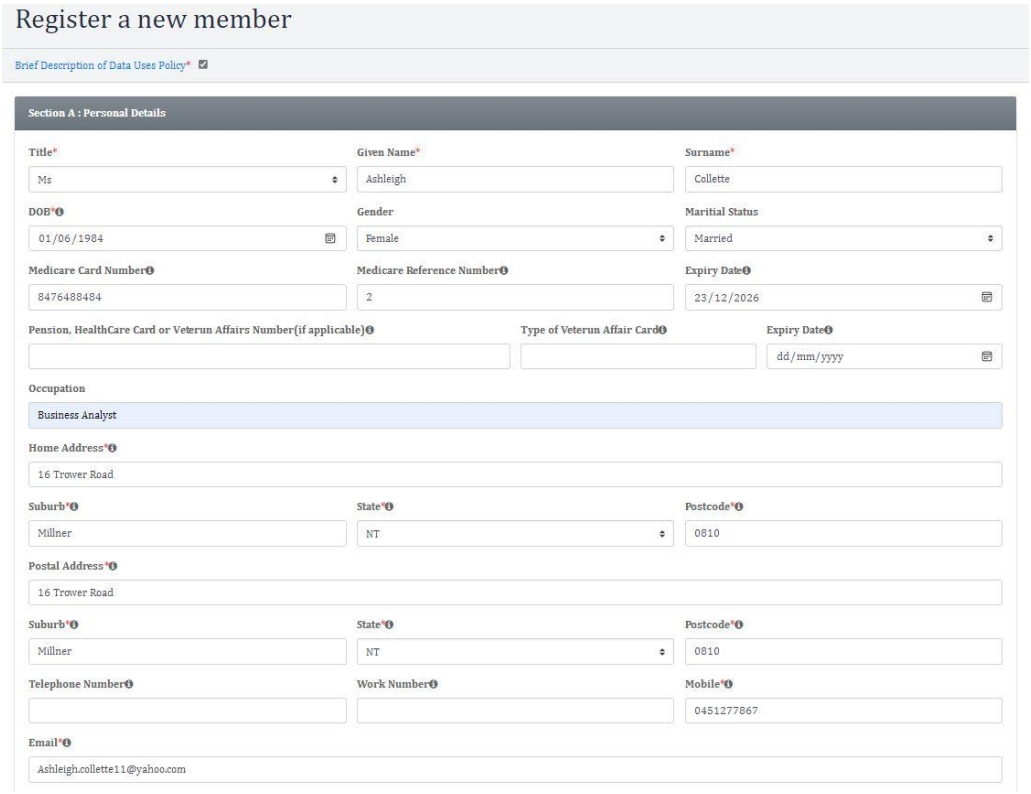


Figure 10. Patient registration application—Section A (Personal Details). Note: Data in the prototype are not original and were generated for testing the application.

Results Messages

	ID	Title	SurName	GivenName	DateOfBirth	Gender	MaritalStatus	MedicareCardNumber
1	050325...	Mrs	Semantha	Farida Habib	1980-01-31	Female	Mrs	123445
2	F1B84...	Ms	Collette	Ashleigh	1984-06-01	Female	Mrs	8476488484

Figure 11. MySQL results—windows authentication—unmasked data.

Results Messages

	ID	Title	SurName	GivenName	DateOfBirth	Gender	MaritalStatus	MedicareCardNumber
1	050325...	xxxx	xxxx	xxxx	xxxx	xxxx	xxxx	1XXX45
2	F1B848...	xxxx	xxxx	xxxx	xxxx	xxxx	xxxx	8XXX84

Figure 12. MySQL results—SQL server authentication—masked data.

IsConsentYes	IsConfirmAgreement	IHI_ID
1	1	120826
1	1	122713

Figure 13. SQL results—Patient ID (IHI_ID) creation.

4.1.2. Patient Registration Prototype (Allergies and Medicines Details)

Section C: Allergies and Medicines Details from the patient registration was implemented in ASP.NET, presented in Figure 14, with a few of the associated unmasked data in windows authentication shown in Figures 15 and 16 and masked data in the SQL Server authentication presented in Figures 17 and 18:

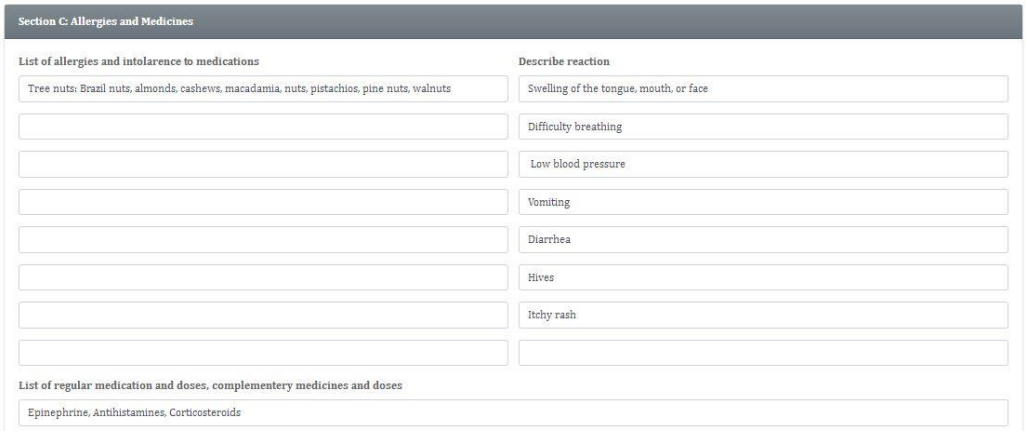


Figure 14. Patient registration—Section C (Allergies and Medicines Details).

Results Messages

	ListOfRegularMedication1
1	Antihistamines,
2	Epinephrine, Antihistamines, Corticosteroids

Figure 15. SQL results—windows authentication—unmasked data.

Results Messages

	ListOfAllergiesAndIntolarence1	ListOfAllergiesAndIntolarenceReaction1
1	xxxx	xxxx
2	xxxx	xxxx

Figure 16. SQL results—windows authentication—unmasked data.

Results Messages

	ListOfAllergiesAndIntolarence1	ListOfAllergiesAndIntolarenceReaction1
1	Cow's Milk	Rashes of the tongue, mouth, or face
2	Tree nuts: Brazil nuts, almonds, cashews, macada...	Swelling of the tongue, mouth, or face

Figure 17. SQL results —SQL server authentication—masked data.

Results Messages

	ListOfRegularMedication1
1	xxxx
2	xxxx

Figure 18. SQL results —SQL server authentication—masked data.

4.1.3. Doctor Registration Prototype (Personal Details)

The Healthcare Provider Identifier—Individual (Doctor) portal was designed using the application framework ASP.NET and SQL server. For example, Personal Details from the doctor registration are presented in Figure 19, with their associated unmasked data in windows authentication in Figure 20 and masked data in the SQL Server authentication in Figure 21.

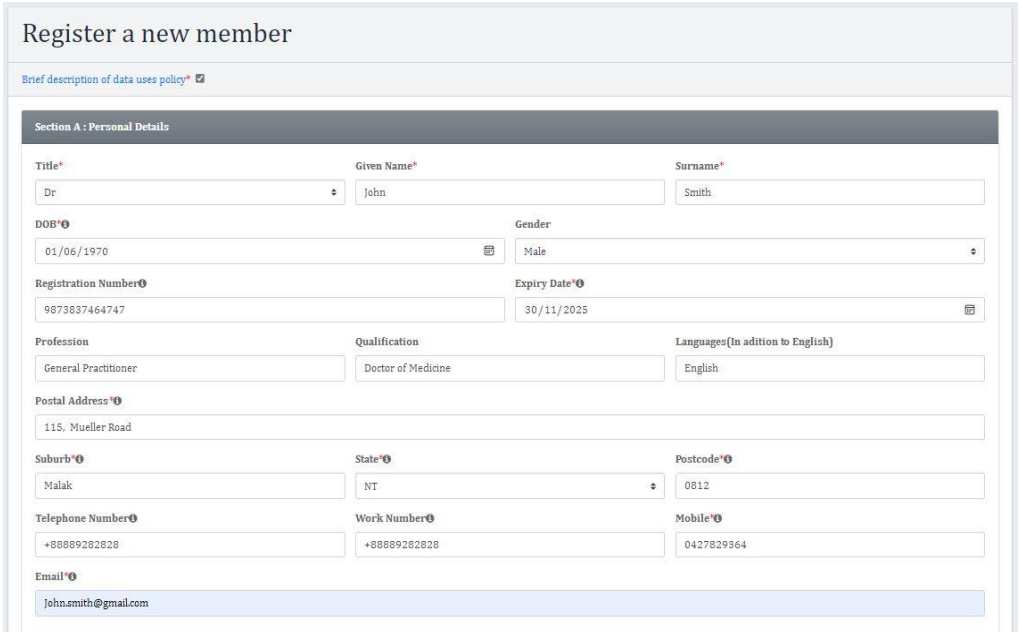


Figure 19. Doctor registration (Section A: Personal Details). Note: Data in the prototype are not original and were generated for testing the application.

Results Messages

	ID	ApplicationUserID	Title	SurName	GivenName	DateOfBirth	Gender	RegistrationNumber
1	9C0A68B...	63797BD8-41D9...	Dr	Roberts	Andrea	1985-09-30	Female	822334346
2	B5906E15...	165F8969-911C-...	Dr	Smith	John	1955-05-01	Male	123456789

Figure 20. SQL results—windows authentication—unmasked data.

Results Messages

	ID	ApplicationUserID	Title	SurName	GivenName	DateOfBirth	Gender	RegistrationNumber
1	9C0A68B...	63797BD8-41D9...	xxxx	xxxx	xxxx	xxxx	xxxx	8XXX46
2	B5906E15...	165F8969-911C-...	xxxx	xxxx	xxxx	xxxx	xxxx	1XXX89

Figure 21. SQL results—SQL server authentication—masked data.

4.1.4. Doctor's Profile Prototype (Patient's Records)

The doctor sends the request, and the patient's EHR retrieves using the Patient ID (IHI_ID), as presented in Figure 22. Additionally, the associated unmasked data in windows authentication are retrieved in Figure 23 and masked data in the SQL Server authentication can be retrieved based on the IHI, as presented in Figure 24.

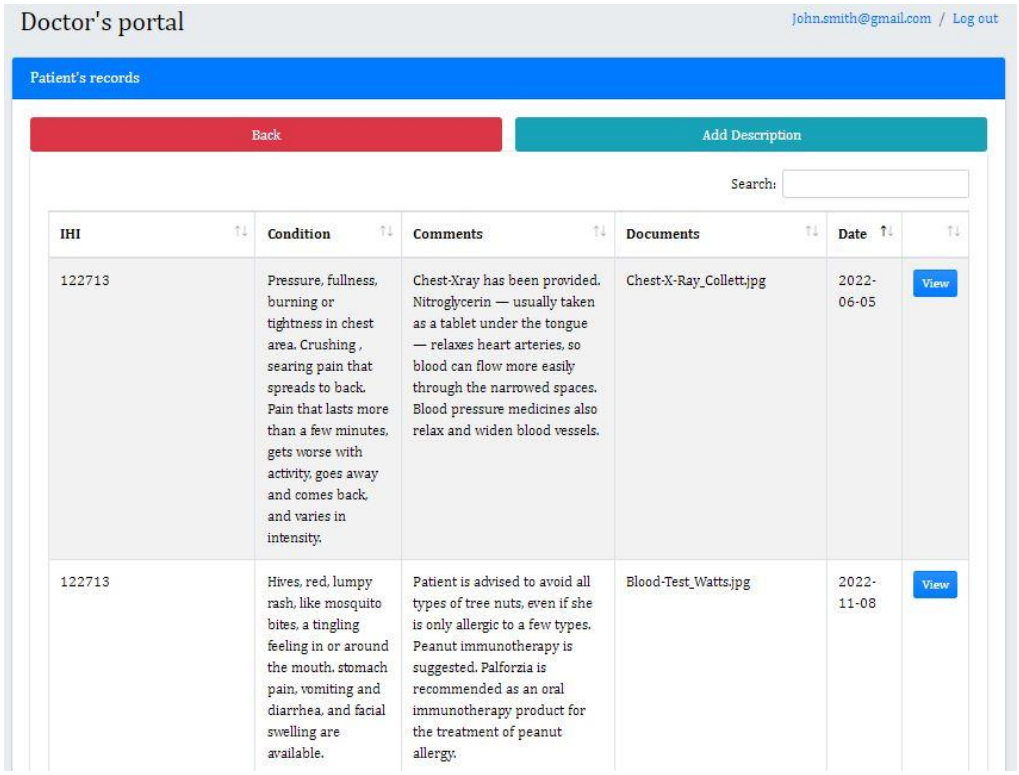


Figure 22. Doctor's portal—patient's records. Note: Data included in the prototype are not original and were generated for testing the application.

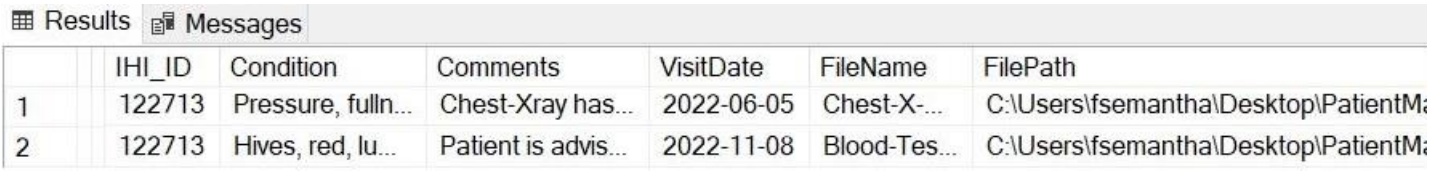


Figure 23. SQL results—windows authentication—unmasked data.

Figure 24. SQL results—SQL server authentication—masked data.

4.1.5. Doctor's Profile Prototype (Adding Records to Patient's Profile)

The process for uploading medical files to the patient's records is shown in Figure 25. Data uploaded to the patients' records are presented in Figure 26. The SQL Server Tables for the uploaded file and file path are shown in Figure 27. Based on the patient ID (IHI), the files are uploaded to the patient's records.

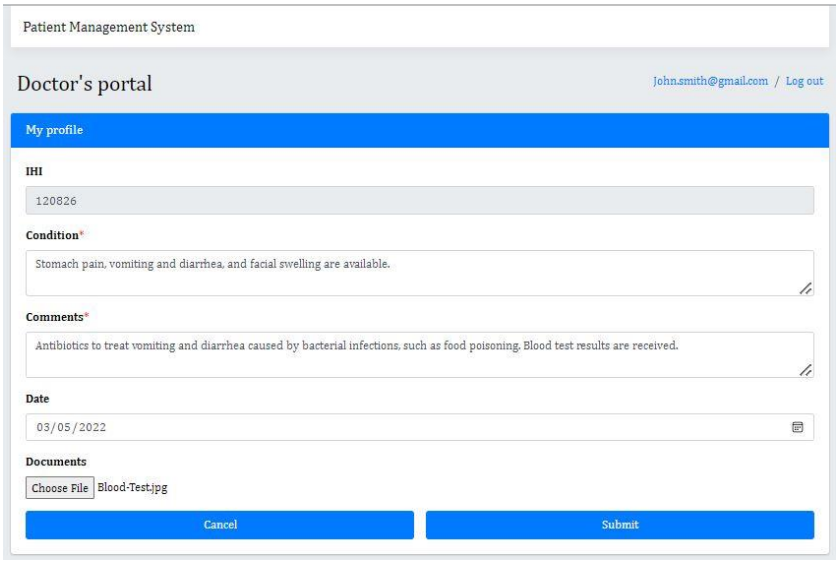
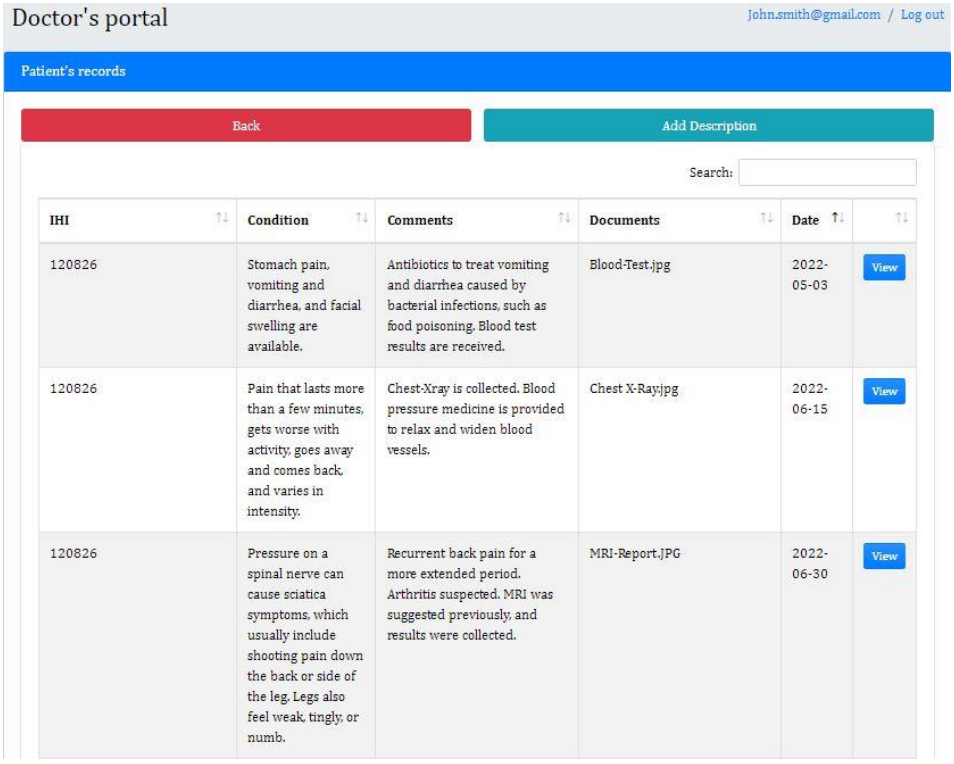


Figure 25. Doctor's Portal—file upload to patient's records.



IHI	Condition	Comments	Documents	Date	
120826	Stomach pain, vomiting and diarrhea, and facial swelling are available.	Antibiotics to treat vomiting and diarrhea caused by bacterial infections, such as food poisoning. Blood test results are received.	Blood-Test.jpg	2022-05-03	View
120826	Pain that lasts more than a few minutes, gets worse with activity, goes away and comes back, and varies in intensity.	Chest-Xray is collected. Blood pressure medicine is provided to relax and widen blood vessels.	Chest X-Ray.jpg	2022-06-15	View
120826	Pressure on a spinal nerve can cause sciatica symptoms, which usually include shooting pain down the back or side of the leg. Legs also feel weak, tingly, or numb.	Recurrent back pain for a more extended period. Arthritis suspected. MRI was suggested previously, and results were collected.	MRI-Report.JPG	2022-06-30	View

Figure 26. Doctor's Portal—data uploaded to patient records.

Results Messages

	HPI-I_ID	IHI_ID	Condition	Comments	VisitDate	FileName	FilePath
1	05032503-...	120826	Pain that lasts m...	Chest-Xray is collect...	2022-06-15	Chest X-Ray...	C:\Users\fsemantha\Desktop\PatientManagementSyste...
2	05032503-...	120826	Stomach pain, v...	Antibiotics to treat vo...	2022-05-03	Blood-Test.jpg	C:\Users\fsemantha\Desktop\PatientManagementSyste...
3	05032503-...	120826	Pressure on a sp...	Recurrent back pain ...	2022-06-30	MRI-Report....	C:\Users\fsemantha\Desktop\PatientManagementSyste...
4	F1B848DB-...	122713	Hives, welts or w...	They are advised to a...	2022-12-20	Blood-Test_...	C:\Users\fsemantha\Desktop\PatientManagementSyste...
5	F1B848DB-...	122713	Hives, red, lumpy...	Patient is advised to ...	2022-11-08	Blood-Test_...	C:\Users\fsemantha\Desktop\PatientManagementSyste...
6	F1B848DB-...	122713	Pressure, fullnes...	Chest-Xray has been...	2022-06-05	Chest-X-Ray...	C:\Users\fsemantha\Desktop\PatientManagementSyste...

Figure 27. Uploaded patient records in the SQL server.

The prototypes showed that the proposed privacy by design components are implemented in the applications for the patients' and the doctors' healthcare record management. While collecting personal and sensitive information, the proposed healthcare principles are applied correspondingly. The collected data are stored in the SQL Table with appropriate masking functions based on their data types. However, the users who have the authentication can access and view the data in unmasked condition. The proposed mechanisms were implemented and presented in the prototypes to confirm all probable consequences to establish the solution. Usability and security testing was conducted to measure the outcome and effectiveness of the proposed framework in the following section.

5. Testing

This section provides the usability and security testing for developing the proposed 'PbDinEHR' prototype.

5.1. Usability Testing

We established and validated the user performance and address potential design concerns to improve the efficiency and end-user satisfaction for the proposed 'PbDinEHR'. We used two tools to carry out the usability testing [106,107]: First Click Testing [109] and Five Seconds Tests [110].

5.1.1. First Click Testing

In First Click Testing, 15 participants explored what they clicked on first on the patient registration interface to read the "Brief description of the data uses in the policy". We used the PRMS patient registration prototype to carry out this test. This testing allows us to evaluate the effectiveness of the proposed prototype to find out the user's response to the navigation and how the users complete their intended tasks [109,110].

➢ *First Click Testing Results*

The heatmap shows the results: "Where would you click to read the brief description of the data uses policy?" As we can see, the majority of the users clicked the link on the top left of the patient registration interface to read the data uses policy, as shown in Figure 28.

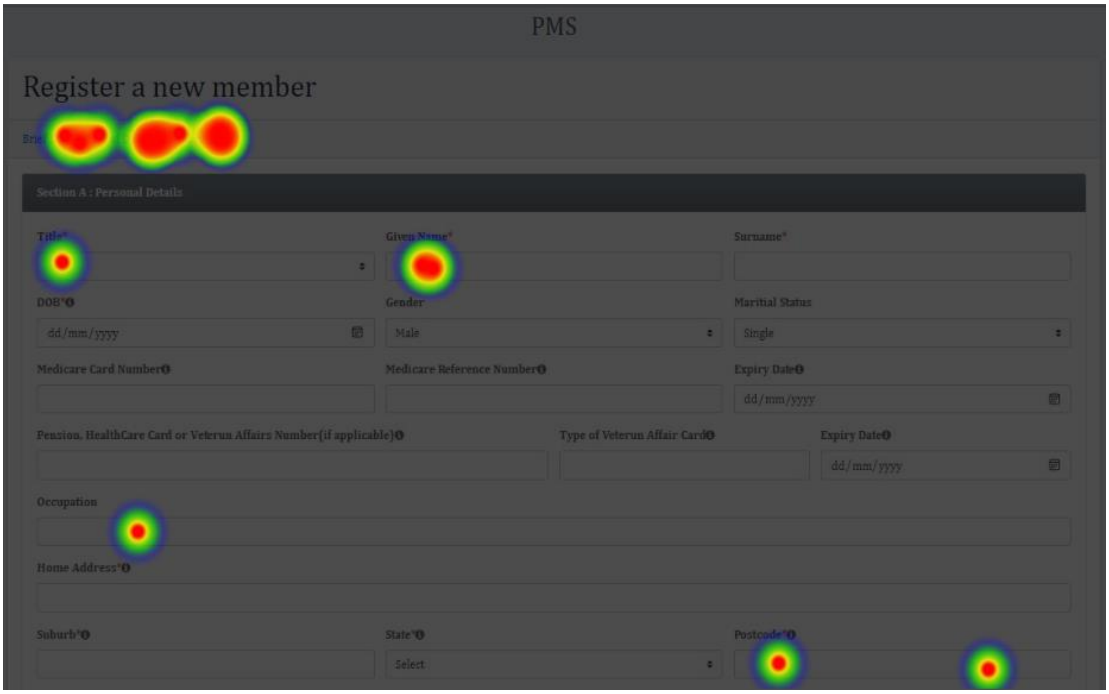


Figure 28. Heatmap results for first click testing.

a. Linear Scale Question

In Figure 29, the linear scale question shows that 73% of users rated excellent for the question, "How would you rate this website in terms of clarity of use?"

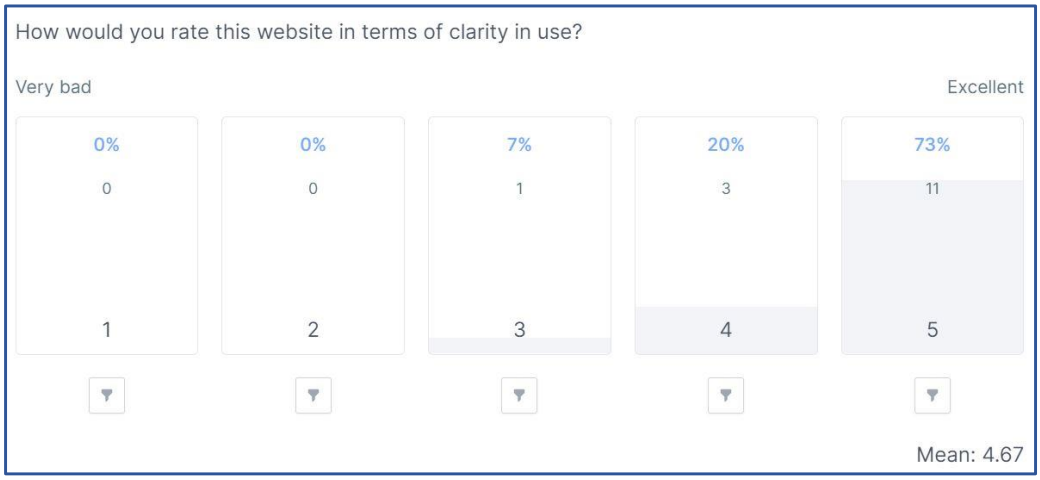


Figure 29. Results—linear scale question.

b. Single Choice Question

Fifteen participants answered the single choice question "What do you think the website is for"; 80% of participants answered "for patients to do registration", 13% of participants answered "for doctors to add a description to patient's profile" and 7% participants selected "other", as presented in Figure 30.

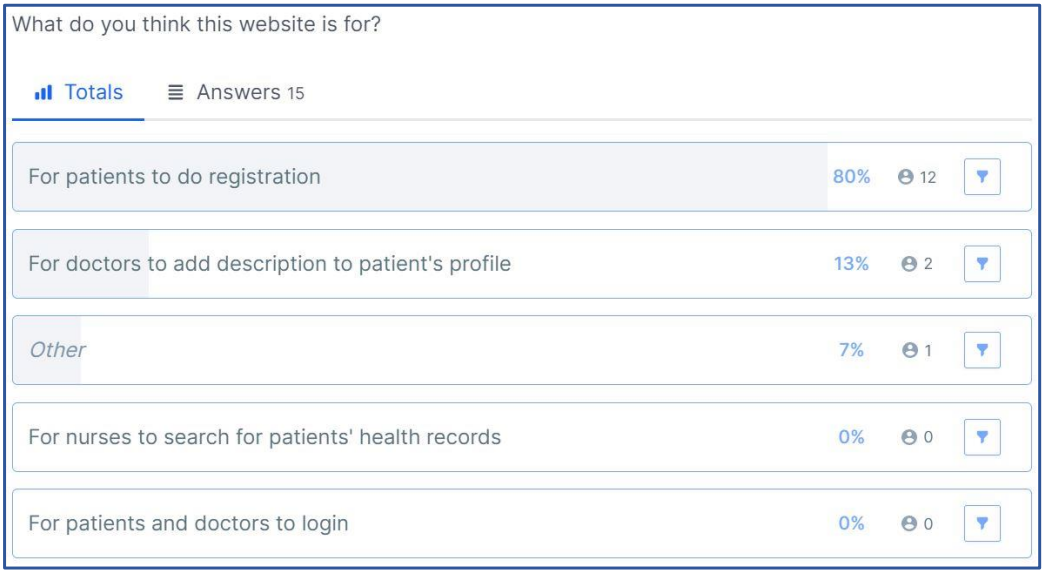


Figure 30. Results—single choice question.

5.1.2. Five Seconds Test

In this method, five seconds were given to 15 participants to view the interface to measure the impression given and the information take away for the users. The participants were given a primer on the format and prompted to pay close attention to the design. This testing determines whether the first impressions of the interface are on point [110].

➢ *Five Seconds Test Results*

The participants were shown a screenshot of the patient registration interface (Figure 31) for five seconds. The participants were asked to answer short and linear scale questions.

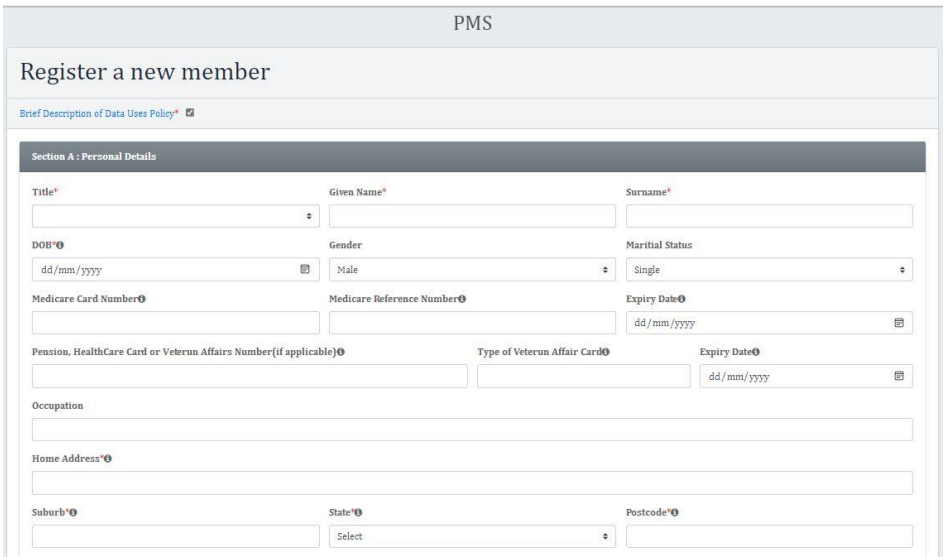


Figure 31. Design shown for five seconds.

a. Word Cloud—Short Text Question

The results for the short text question "What is this website for" are shown in terms of a word cloud in Figure 32.

Figure 32. Results—word cloud—short text question.

b. Linear Scale Question

In Figure 33, the linear scale question shows that 75% of users rated excellent for the question, "How would you rate this website in terms of design?"

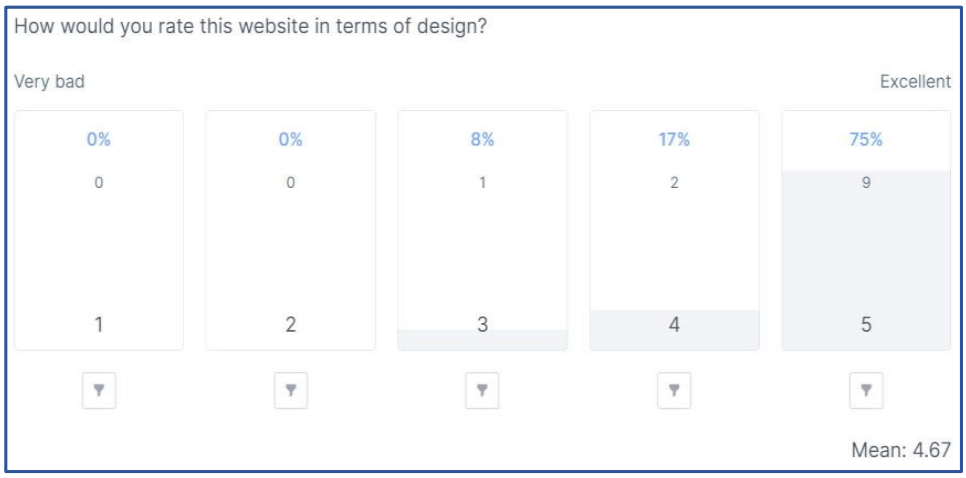


Figure 33. Results—linear scale question.

5.2. Security Testing

Security testing was conducted using Zed Attack Proxy (ZAP) tool, an open source web application security scanner that identifies the security vulnerabilities of the proposed PRMS prototype and provides possible solutions based on the identified alerts categories [111].

➢ Sites

Our PRMS prototype runs on localhost, and the following sites were included for testing: https://localhost:44355 (accessed on 12 February 2023).

The security alerts of our PRMS prototype were identified using the ZAP tool and listed in Figure 34, which shows three medium, five low and five informational alerts; however, no high priority alerts were involved.

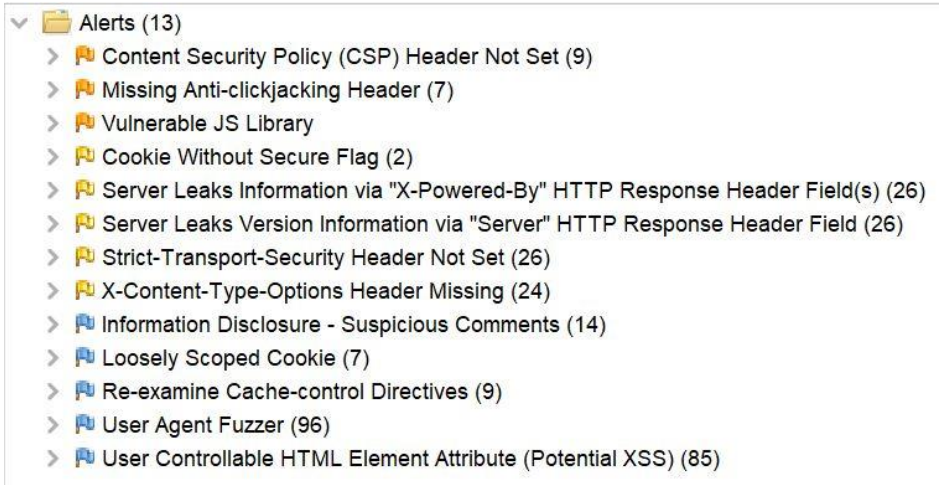


Figure 34. Security alerts in ZAP tool.

➢ **Risk Levels**

High, Medium, Low, Informational

➢ **Confidence Levels**

User Confirmed, High, Medium, Low

Figure 35 shows the number of alerts for each level of risk and confidence included in the report. The percentages in brackets represent the count as a percentage of the total number of alerts included in the report, rounded to one decimal place. Indeed, the user confirmed that there were no high priority alerts involved. The overall significances of our research findings are presented in the following section.

		Confidence				
		User Confirmed	High	Medium	Low	Total
Risk	High	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
	Medium	0 (0.0%)	1 (7.7%)	2 (15.4%)	0 (0.0%)	3 (23.1%)
	Low	0 (0.0%)	2 (15.4%)	3 (23.1%)	0 (0.0%)	5 (38.5%)
	Informational	0 (0.0%)	0 (0.0%)	1 (7.7%)	4 (30.8%)	5 (38.5%)
	Total	0 (0.0%)	3 (23.1%)	6 (46.2%)	4 (30.8%)	13 (100%)

Figure 35. Alerts count by risks and confidence.

6. Discussion

In paper [21], we conducted a systematic literature review to extensively evaluate privacy by design key contexts to identify the limitations of existing frameworks. Based on the limitations specified in the review paper, we extended the research and proposed a conceptual framework in paper [22] that ensures privacy in the patient record management system. The relevant research studies primarily explored and examined privacy by design studies to identify the vital mechanisms critical to designing a comprehensive privacy-preserving solution (see Section 2).

This research proposes a holistic framework that incorporates the consequences of prior study and extension with modern and globally verified fundamental mechanisms that ensure maximum privacy in every layer of data processing. 'PbDinEHR' is not a single component, but a collaboration of data privacy components that are internationally recognised (see Section 3). Individual data privacy components are distinctly assessed to incorporate into the proposed framework to establish a scalable and secure system for patients' personal and sensitive information management. Healthcare principles (HPs) are developed by analysing the existing privacy by design principles. The compliance between the proposed HPs is assessed with the necessary aspects to verify the reliability of the principles in healthcare systems. Dynamic Data Masking (DDM) helps to establish parameters for the data types based on sensitivity to limit sensitive data exposure. In addition, Transparent Database Encryption (TDE) sets up access control and permissions with valid credentials to prevent unauthorized access towards healthcare records. Moreover, Privacy Impact Assessment (PIA) establishes the valuation whereby all the nominated healthcare principles are assessed to check if they comply with the standards or not. To do so, compatibilities of the proposed HPs are established with internationally verified standards, APPs and GDPR.

In this paper, we collaborated on all these core components to design the proposed framework. Based on the research outcome, a healthcare application is developed using ASP.NET and SQL Server, which guarantees maximum privacy preservation in accessing and retrieving patients' healthcare records (see Section 4). In addition, to enhance the scalability of the proposed framework, IPFS and blockchain are used for sharing medical files and recording the transactions where patients, doctors, hospitals, and other providers share patient-centric data securely and transparently in a distributed environment. We conducted usability and security testing to ensure its effectiveness and security (See Section 5). Our research identified the gaps, designed the framework, validated framework components with standards, and developed prototypes. The proposed framework 'PbDinEHR' provides privacy measurements that can be embedded to ensure data preservation from the beginning of healthcare data management.

7. Conclusions and Future Works

As the world is moving towards the new information age, governing data breaches and ensuring information privacy play a crucial role. Traditional patient electronic health record (EHR) systems are complex, expensive, centralized, and often insecurely store and share patients' healthcare data. Furthermore, research into data privacy by design for healthcare records is relatively behind as the confidentiality of patients' EHR is not prioritized in many systems. As new technology matures, researchers better comprehend the intricacies of data privacy and security technologies to design EHR systems in innovative ways. Our research designed and implemented a comprehensive privacy by design solution for healthcare systems with an accumulation of internationally verified components: privacy by design fundamental principles, privacy design strategies, standards, privacy impact assessment, data decentralisation, and distributed file system technologies. The proposed framework was designed with systematic activity through three distinct phases of system design, including planning, assessment, and implementation. The purpose of this framework was to integrate essential data privacy by design mechanisms into one place while collecting, managing and storing personal information; therefore, the healthcare system can ensure the maximum privacy of the patient's personal data. In the last few years, blockchain has offered advancements to prototype, simulate and launch custom blockchain network implementations that are easy to apply to medical records while sharing information in a distributed environment. The permissioned blockchain networks are used to hold actions related to file requests and response processes. Aside from blockchain, IPFS has also gained popularity in the expedition of EHR enhancement. To securely store the medical files of different healthcare organisations, IPFS provides decentralised and distributed file storage. IPFS supports against malicious fighting attacks by ensuring no single point of failure for storing valuable medical data.

In conclusion, privacy by design can be applied to a variety of domains and industries to ensure that privacy is considered in the design process from the initiation of any system, rather than treating privacy an afterthought. Organisations can protect the privacy of their users and can build trust with their users by applying privacy by design mechanisms. Privacy by design frameworks can be utilized in a variety of domains—software development, financial services, government, education, marketing, etc.—while dealing with personal and sensitive user data. In our proposed framework, we applied privacy by design principles, privacy design strategies, and a privacy impact assessment in each layer of healthcare data processing. The main benefit associated with our work is that this research has generalizability for other data sharing domains. However, this may fall short of requirements from other domains. This is one of the limitations of our framework. On the other hand, the immutability of blockchain technology and the GDPR's right to erasure can appear to conflict with each other. In this research, our framework is compatible with GDPR and blockchain is applied for sharing medical files and recording the transactions to ensure scalability; thus, all the personal and sensitive data records are not designed to be decentralised. Therefore, in our proposed framework, blockchain technology was carefully

considered to balance the immutability of the blockchain with the GDPR's right to erasure, which is a benefit to this research.

In our future endeavours, firstly, we will extend our work to other data sharing domains that manage sensitive user data and regulatory frameworks. Secondly, we will conduct more research on the immutability of blockchain technology and the data privacy standard's right to erasure in the case of any conflict with each other. Thirdly, we will carry out analysis on more robust encryption techniques to enhance the selected encryption techniques for better consequences in protecting patient data. Homomorphic encryption will be applied that will allow complex operations on encrypted data without compromising the encryption. In addition, secure multi-party computation will be considered for privacy, preserving computation which is a cryptography subfield to equally compute functions while keeping the inputs private [112]. Fourthly, we intend to extend the database users and access control, as only limited user levels are assigned to check the functionality of the proposed framework. Fifthly, three healthcare provider nodes will be used to design a private IPFS network that securely stores medical records. Therefore, more IPFS nodes will be included in the network for medical file sharing. Finally, the proposed application uses the user's email to log in and access the healthcare records. Mobile based access will be established for more reliable user control over healthcare records management. Our proposed framework allows healthcare systems users to manage their medical records and ensure maximum privacy while sharing them with various healthcare providers. The resulting framework delivers advanced incorporation and allocation of personal data with maximum integrity and confidentiality of the healthcare system to decrease data breaches worldwide.

Author Contributions: Conceptualization, F.H.S. and S.A.; methodology, F.H.S. and S.A.; software, F.H.S. and S.A.; validation, F.H.S. and B.S.; formal analysis, F.H.S. and K.C.Y.; investigation, F.H.S. and S.A.; resources, B.S.; data curation, F.H.S.; writing—original draft preparation, F.H.S. and S.A.; writing—review and editing, F.H.S.; visualization, F.H.S. and S.A.; supervision, S.A. and B.S.; project administration, S.A. and K.C.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Not applicable.

Acknowledgments: The authors acknowledge the support of the Faculty of Science and Technology, Charles Darwin University.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Appendix A.1. Query to Assign Access Controls to Database Users

```
CREATE ROLE ACCESSCONTROL1
GRANT SELECT, ALTER, INSERT, UPDATE,DELETE ON PMS.dbo.ApplicationUsers TO ACCESSCONTROL1
GRANT SELECT, ALTER, INSERT, UPDATE,DELETE ON PMS.dbo.UserAccess TO ACCESSCONTROL1
CREATE ROLE ACCESSCONTROL2
GRANT SELECT, ALTER, INSERT, UPDATE, DELETE ON PMS.dbo.PatientRecords TO ACCESSCONTROL2
GRANT SELECT, ALTER, INSERT, UPDATE, DELETE ON PMS.dbo.Doctors TO ACCESSCONTROL2
GRANT SELECT, ALTER, INSERT, UPDATE, DELETE ON PMS.dbo.Patients TO ACCESSCONTROL2
CREATE ROLE ACCESSCONTROL3
GRANT SELECT, INSERT ON PMS.dbo.PatientRecords TO ACCESSCONTROL3
```

Figure A1. Assigning different access controls to database users.

Appendix A.2.

Table A1. Root Node Address to the Bootstrap IPFS Nodes.

Root Node Swarmed Peers
• /ip4/<node 2 IP>/tcp/<port address>/ipfs/<Peer ID of node 2> • /ip4/<node 3 IP>/tcp/<port address>/ipfs/<Peer ID of node 3> • /ip4/<node 4 IP>/tcp/<port address>/ipfs/<Peer ID of node 4>
Node 2 Swarmed Peers
• /ip4/<root node IP>/tcp/<port address>/ipfs/<Peer ID of root node>
Node 3 Swarmed Peers
• /ip4/<root node IP>/tcp/<port address>/ipfs/<Peer ID of root node>
Node 4 Swarmed Peers
• /ip4/<root node IP>/tcp/<port address>/ipfs/<Peer ID of root node>

Appendix A.3. Smart Contract Transactions Tracking on Blockchain

```
pragma solidity ^0.5.10;
contract Request {
        // Address represents an external (user) account.
    address public owner;
        // A `mapping` is essentially a hash table data structure.
    mapping (address => uint) public balances;
        // Events allow for logging of activity on the blockchain.
        // Ethereum clients can listen for events in order to react with change in state.
    event Transfer(address from, address to, uint amount);
        // to the address of the contract creator.
    constructor() public {
        // All smart contracts rely on external transactions to trigger its       functions.
        // `msg` is a global variable that includes relevant data on the given    transaction, such as
the IHI (individual healthcare identifier) of a patient.
        owner = msg.sender;
         }
```

Figure A2. Smart contract to track transactions on blockchain.

Appendix A.4. Request/Response Transaction Object

```
{
  from: "0xEA674fdDe714fd979de3EdF0F56AA9716B898ec8",
  to: "0xac03bb73b6a9e108530aff4df5077c2b3d481e5a",
  gasLimit: "21000",
  maxFeePerGas: "300",
  maxPriorityFeePerGas: "10",
  nonce: "0",
  value: "10000000000"
}
```

Figure A3. A request/response transaction object.

References

1. Martin, K.D.; Murphy, P.E. The role of data privacy in marketing. *J. Acad. Mark. Sci.* **2017**, *45*, 135–155. [CrossRef]
2. Avery, A. After the disclosure: Measuring the short-term and long-term impacts of data breach disclosures on the financial performance of organizations. *Inf. Comput. Secur.* **2021**, *29*, 500–525. [CrossRef]
3. Gwebu, K.L.; Wang, J.; Wang, L. The role of corporate reputation and crisis response strategies in data breach management. *J. Manag. Inf. Syst.* **2018**, *35*, 683–714. [CrossRef]

4. Powell, O. The Biggest Data Breaches and Leaks of 2022. Available online: https://www.cshub.com/attacks/articles/the-biggest-data-breaches-and-leaks-of-2022 (accessed on 15 January 2023).
5. Kovacs, E. Over 50,000 Revolut Customers Affected by Data Breach. Available online: https://www.securityweek.com/over-50000-revolut-customers-affected-data-breach/ (accessed on 15 January 2023).
6. Lauver, M. Data Breach Exposes Records of 2.5 Million Student Loan Borrowers. Available online: https://www.securitymagazine.com/articles/98306-data-breach-exposes-records-of-25-million-student-loan-borrowers (accessed on 15 January 2023).
7. Brown, H. Privacy law and cyber security: Is your practice secure?: Client confidentiality and data breach. *Law Soc. J.* **2017**, 88–89.
8. Commission, A.-A.S.I. Guidance for Consumers Impacted by the Optus Data Breach. Available online: https://asic.gov.au/about-asic/news-centre/news-items/guidance-for-consumers-impacted-by-the-optus-data-breach/ (accessed on 20 December 2022).
9. Kruger, C. AFP steps in as Medibank hack data migrates from dark web. *The Sydney Morning Herald*, 15 November 2022.
10. Barbaschow, A. Medibank Hackers Declare the 'Case Closed' as Privacy Commissioner Launches Investigation. *Gizmodo Australia*, 2 December 2022.
11. Cubby, B. 130,000 Telstra customers exposed in data breach. *The Sydney Morning Herald*, 10 December 2022.
12. Seh, A.H.; Zarour, M.; Alenezi, M.; Sarkar, A.K.; Agrawal, A.; Kumar, R.; Ahmad Khan, R. Healthcare data breaches: Insights and implications. *Healthcare* **2020**, *8*, 133. [CrossRef]
13. LaMonica, H.M.; Roberts, A.E.; Lee, G.Y.; Davenport, T.A.; Hickie, I.B. Privacy Practices of Health Information Technologies: Privacy Policy Risk Assessment Study and Proposed Guidelines. *J. Med Internet Res.* **2021**, *23*, e26317. [CrossRef]
14. El Ouazzani, Z.; El Bakkali, H.; Sadki, S. Privacy Preserving in Digital Health: Main Issues, Technologies, and Solutions. In *Research Anthology on Privatizing and Securing Data*; IGI Global: Hershey, PA, USA, 2021; pp. 1503–1526.
15. Chenthara, S.; Ahmed, K.; Wang, H.; Whittaker, F. Security and privacy-preserving challenges of e-health solutions in cloud computing. *IEEE Access* **2019**, *7*, 74361–74382. [CrossRef]
16. Hathaliya, J.J.; Tanwar, S. An exhaustive survey on security and privacy issues in Healthcare 4.0. *Comput. Commun.* **2020**, *153*, 311–335. [CrossRef]
17. OAIC. Australian Privacy Principles. Available online: https://www.oaic.gov.au/privacy/australian-privacy-principles/ (accessed on 5 July 2022).
18. Tamburri, D.A. Design principles for the General Data Protection Regulation (GDPR): A formal concept analysis and its evaluation. *Inf. Syst.* **2019**, *91*, 101469. [CrossRef]
19. Tauqeer, A.; Kurteva, A.; Chhetri, T.R.; Ahmeti, A.; Fensel, A. Automated GDPR Contract Compliance Verification Using Knowledge Graphs. *Information* **2022**, *13*, 447. [CrossRef]
20. Chhetri, T.R.; Kurteva, A.; DeLong, R.J.; Hilscher, R.; Korte, K.; Fensel, A. Data Protection by Design Tool for Automated GDPR Compliance Verification Based on Semantically Modeled Informed Consent. *Sensors* **2022**, *22*, 2763. [CrossRef]
21. Semantha, F.H.; Azam, S.; Yeo, K.C.; Shanmugam, B. A systematic literature review on privacy by design in the healthcare sector. *Electronics* **2020**, *9*, 452. [CrossRef]
22. Semantha, F.H.; Azam, S.; Shanmugam, B.; Yeo, K.C.; Beeravolu, A.R. A Conceptual Framework to Ensure Privacy in Patient Record Management System. *IEEE Access* **2021**, *9*, 165667–165689. [CrossRef]
23. OVIC. *Privacy by Design: Effective Privacy Management in the Victorian Public Sector*; Office of the Victorian Information Commissioner: Melbourne, Australia, 2019; pp. 1–8.
24. OVIC. Privacy Impact Assessment Guide. Available online: https://ovic.vic.gov.au/privacy/for-agencies/privacy-impact-assessments/ (accessed on 10 January 2023).
25. Moncrieff, S.; Venkatesh, S.; West, G. A framework for the design of privacy preserving pervasive healthcare. In Proceedings of the 2009 IEEE International Conference on Multimedia and Expo, New York, NY, USA, 28 June–3 July 2009; pp. 1696–1699.
26. Tariq, F.; Khan, Z.; Sultana, T.; Rehman, M.; Shahzad, Q.; Javaid, N. *Leveraging Fine-grained Access Control in Blockchain-Based Healthcare System*; Advances in Intelligent Systems and Computing; Springer: Berlin/Heidelberg, Germany, 2020.
27. Nagasubramanian, G.; Sakthivel, R.K.; Patan, R.; Gandomi, A.H.; Sankayya, M.; Balusamy, B. Securing e-health records using keyless signature infrastructure blockchain technology in the cloud. *Neural Comput. Appl.* **2020**, *32*, 639–647. [CrossRef]
28. Thwin, T.; Vasupongayya, S. Blockchain-Based Access Control Model to Preserve Privacy for Personal Health Record Systems. *Secur. Commun. Networks* **2019**, *2019*, 8315614. [CrossRef]
29. Wang, H.; Song, Y. Secure Cloud-Based EHR System Using Attribute-Based Cryptosystem and Blockchain. *J. Med Syst.* **2018**, *42*, 152. [CrossRef]
30. Roda, C.; Kennedy, B.; Perry, S.; del Álamo, M.; Tsormpatzoudi, P.; Coudert, F.; Elshaafi, H.; Kargl, F.; Kopp, H. PReparing Industry to Privacy-by-design by supporting its Application in REsearch. Available online: https://ac.aup.edu/~croda/publications/PRIPARE_Deliverable_D4.1_v1.pdf (accessed on 12 February 2023).
31. Miyachi, K.; Mackey, T.K. hOCBS: A privacy-preserving blockchain framework for healthcare data leveraging an on-chain and off-chain system design. *Inf. Process. Manag.* **2021**, *58*, 102535. [CrossRef]
32. Shrestha, N.; Alsadoon, A.; Prasad, P.; Hourany, L.; Elchouemi, A. Enhanced e-health framework for security and privacy in healthcare system. In Proceedings of the 2016 6th International Conference on Digital Information Processing and Communications (ICDIPC), Beirut, Lebanon, 21–23 April 2016; pp. 75–79.

33. Bhattacharya, P.; Tanwar, S.; Bodke, U.; Tyagi, S.; Kumar, N. BinDaaS: Blockchain-Based Deep-Learning as-a-Service in Healthcare 4.0 Applications. *IEEE Trans. Netw. Sci. Eng.* **2019**, *8*, 1242–1255. [CrossRef]
34. Huang, J.; Qi, Y.W.; Asghar, M.R.; Meads, A.; Tu, Y. MedBloc: A Blockchain-Based Secure EHR System for Sharing and Accessing Medical Data. In Proceedings of the 2019 18th IEEE International Conference on Trust, Security and Privacy in Computing and Communications/13th IEEE International Conference on Big Data Science and Engineering (TrustCom/BigDataSE), Rotorua, New Zealand, 5–8 August 2019; pp. 594–601.
35. Perera, C.; McCormick, C.; Bandara, A.K.; Price, B.A.; Nuseibeh, B. Privacy-by-design framework for assessing internet of things applications and platforms. In Proceedings of the 6th International Conference on the Internet of Things, Stuttgart, Germany, 7–9 November 2016; pp. 83–92.
36. Abdul-Ghani, H.A.; Konstantas, D. A comprehensive study of security and privacy guidelines, threats, and countermeasures: An IoT perspective. *J. Sens. Actuator Netw.* **2019**, *8*, 22. [CrossRef]
37. Foukia, N.; Billard, D.; Solana, E. PISCES: A framework for privacy by design in IoT. In Proceedings of the 2016 14th Annual Conference on Privacy, Security and Trust (PST), Auckland, New Zealand, 12–14 December 2016; pp. 706–713.
38. Hyla, T.; Pejaś, J. eHealth Integrity Model Based on a Permissioned Blockchain. In Proceedings of the 2019 Cybersecurity and Cyberforensics Conference (CCC), Melbourne, Australia, 8–9 May 2019; pp. 172–177.
39. Morales-Trujillo, M.E.; Garcia-Mireles, G.A. Extending ISO/IEC 29110 basic profile with privacy-by-design approach: A case study in the health care sector. In Proceedings of the 2018 11th International Conference on the Quality of Information and Communications Technology (QUATIC), Coimbra, Portugal, 4–7 September 2018; pp. 56–64.
40. Li, C.T.; Shih, D.H.; Wang, C.C.; Chen, C.L.; Lee, C.C. A Blockchain Based Data Aggregation and Group Authentication Scheme for Electronic Medical System. *IEEE Access* **2020**, *8*, 173904–173917. [CrossRef]
41. Bari, L.; O'Neill, D.P. Rethinking patient data privacy in the era of digital health. *Health Aff.* **2019**, *12*.
42. Zaeem, R.N.; Barber, K.S. The effect of the GDPR on privacy policies: Recent progress and future promise. *ACM Trans. Manag. Inf. Syst. (TMIS)* **2020**, *12*, 1–20. [CrossRef]
43. Shuaib, M.; Alam, S.; Alam, M.S.; Nasir, M.S. Compliance with HIPAA and GDPR in blockchain-based electronic health record. *Mater. Today Proc.* **2021**. [CrossRef]
44. Baik, J.S. Data privacy against innovation or against discrimination?: The case of the California Consumer Privacy Act (CCPA). *Telemat. Inform.* **2020**, *52*, 101431. [CrossRef]
45. Cohen, I.G.; Mello, M.M. HIPAA and protecting health information in the 21st century. *JAMA* **2018**, *320*, 231–232. [CrossRef]
46. McKinstry, C.J. The HIPAA privacy rule: Flawed privacy exposed when compared with the European Union's general data protection regulation. *J. Health Care Financ.* **2018**, *45*, 1.
47. Reen, G.S.; Mohandas, M.; Venkatesan, S. Decentralized Patient Centric e-Health Record Management System using Blockchain and IPFS. In Proceedings of the 2019 IEEE Conference on Information and Communication Technology, Allahabad, India, 6–8 December 2019; pp. 1–7.
48. Chenthara, S.; Ahmed, K.; Wang, H.; Whittaker, F.; Chen, Z. Healthchain: A novel framework on privacy preservation of electronic health records using blockchain technology. *PLoS ONE* **2020**, *15*, e0243043. [CrossRef]
49. Reegu, F.A.; Al-Khateeb, M.O.; Zogaan, W.A.; Al-Mousa, M.R.; Alam, S.; Al-Shourbaji, I. Blockchain-based framework for interoperable electronic health record. *Ann. Rom. Soc. Cell Biol.* **2021**, *25*, 6486–6495.
50. Vishnoi, M. MedFabric4Me: Blockchain Based Patient Centric Electronic Health Records System. Mater's Thesis, Arizona State University, Ann Arbor, MI, USA, 2020.
51. Fatokun, T.; Nag, A.; Sharma, S. Towards a Blockchain Assisted Patient Owned System for Electronic Health Records. *Electronics* **2021**, *10*, 580. [CrossRef]
52. Hussien, H.M.; Yasin, S.M.; Udzir, N.I.; Ninggal, M.I.H. Blockchain-based access control scheme for secure shared personal health records over decentralised storage. *Sensors* **2021**, *21*, 2462. [CrossRef]
53. Liang, X.; Zhao, J.; Shetty, S.; Liu, J.; Li, D. Integrating blockchain for data sharing and collaboration in mobile healthcare applications. In Proceedings of the 2017 IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC), Montreal, QC, Canada, 8–13 October 2017; pp. 1–5. [CrossRef]
54. Demir, O.; Kocak, B. A Decentralized File Sharing Framework for Sensitive Data. In Proceedings of the International Conference on Big Data Innovations and Applications, Istanbul, Turkey, 26–28 August 2019; pp. 142–149.
55. Keshta, I.; Odeh, A. Security and privacy of electronic health records: Concerns and challenges. *Egypt. Inform. J.* **2020**, *22*, 177–183. [CrossRef]
56. Cernian, A.A.-O.; Tiganoaia, B.; Sacala, I.A.-O.; Pavel, A.; Iftemi, A. PatientDataChain: A Blockchain-Based Approach to Integrate Personal Health Records. *Sensors* **2020**, *20*, 6538. [CrossRef]
57. George, J.; Bhila, T. Security, confidentiality and privacy in health of healthcare data. *Int. J. Trend Sci. Res. Dev.* **2019**, *3*, 373–377. [CrossRef]
58. Mahore, V.; Aggarwal, P.; Andola, N.; Raghav; Venkatesan, S. Secure and Privacy Focused Electronic Health Record Management System using Permissioned Blockchain. In Proceedings of the 2019 IEEE Conference on Information and Communication Technology, Allahabad, India, 6–8 December 2019; pp. 1–6.
59. Meier, P.; Beinke, J.H.; Fitte, C.; Schulte to Brinke, J.; Teuteberg, F. Generating design knowledge for blockchain-based access control to personal health records. *Inf. Syst. e-Bus. Manag.* **2021**, *19*, 13–41. [CrossRef]

60. Hylock, R.A.-O.; Zeng, X.A.-O. A Blockchain Framework for Patient-Centered Health Records and Exchange (HealthChain): Evaluation and Proof-of-Concept Study. *J. Med. Internet Res.* **2019**, *21*, e13592. [CrossRef]
61. Roehrs, A.; da Costa, C.A.; da Rosa Righi, R. OmniPHR: A distributed architecture model to integrate personal health records. *J. Biomed. Inform.* **2017**, *71*, 70–81. [CrossRef]
62. Shevkar, S.; Patel, P.; Majumder, S.; Singh, H.; Jaglan, K.; Shalu, H. EMRs with blockchain: A distributed democratised electronic medical record sharing platform. *arXiv* **2020**, arXiv:2012.05141.
63. Cavoukian, A. Understanding How to Implement Privacy by Design, One Step at a Time. *IEEE Consum. Electron. Mag.* **2020**, *9*, 78–82. [CrossRef]
64. Hoepman, J.-H. Privacy design strategies. In Proceedings of the IFIP International Information Security Conference, Marrakech, Morocco, 2–4 June 2014; Springer: Berlin/Heidelberg, Germany, 2014; pp. 446–459.
65. Baranchikov, A.I.; Gromov, A.Y.; Gurov, V.S.; Grinchenko, N.N.; Babaev, S.I. The technique of dynamic data masking in information systems. In Proceedings of the 2016 5th Mediterranean Conference on Embedded Computing (MECO), Bar, Montenegro, 12–16 June 2016; pp. 473–476.
66. Jain, R.B.; Puri, M.; Jain, U. A robust dynamic data masking transformation approach to safeguard sensitive data. *Int. J. Future Revolut. Comput. Sci. Commun. Eng.* **2018**, *4*, 366–370.
67. Microsoft. SQL Server—Dynamic Data Masking. Available online: https://learn.microsoft.com/en-us/sql/relational-databases/security/dynamic-data-masking?view=sql-server-ver16 (accessed on 15 October 2022).
68. Natarajan, K.; Shaik, V. Transparent Data Encryption: Comparative Analysis and Performance Evaluation of Oracle Databases. In Proceedings of the 2020 5th International Conference on Research in Computational Intelligence and Communication Networks (ICRCICN), Bangalore, India, 26–27 November 2020; pp. 137–142.
69. Abdul Rahoof, T.; Deepthi, V. Healthchain: A secure scalable health care data management system using blockchain. In Proceedings of the International Conference on Distributed Computing and Internet Technology, Bhubaneswar, India, 9–12 January 2020; pp. 380–391.
70. Vemou, K.; Karyda, M. Evaluating privacy impact assessment methods: Guidelines and best practice. *Inf. Comput. Secur.* **2019**, *28*, 35–53. [CrossRef]
71. Dorri, A.; Steger, M.; Kanhere, S.S.; Jurdak, R. Blockchain: A distributed solution to automotive security and privacy. *IEEE Commun. Mag.* **2017**, *55*, 119–125. [CrossRef]
72. Kayastha, M.; Karim, S.; Sandu, R.; Gide, E. Ethereum Blockchain and Inter-Planetary File System (IPFS) based Application Model to Record and Share Patient Health Information: An Exemplary Case Study for e-Health Education in Nepal. In Proceedings of the 2021 19th International Conference on Information Technology Based Higher Education and Training (ITHET), Sydney, Australia, 4–6 November 2021; pp. 1–7.
73. Vemou, K.; Karyda, M. An Evaluation Framework for Privacy Impact Assessment Methods. In Proceedings of the 12th Mediterranean Conference on Information Systems (MCIS), Corfu, Greece, 28–30 September 2018; p. 5.
74. Fillmore, A.R.; McKinley, C.D.; Tallman, E.F. Managing privacy, confidentiality, and risk: Towards trust. In *Health Information Exchange*; Elsevier: Amsterdam, The Netherlands, 2023; pp. 131–147.
75. Xue, R.; Baron, C.; Esteban, P. Optimising product development in industry by alignment of the ISO/IEC 15288 systems engineering standard and the PMBoK guide. *Int. J. Prod. Dev.* **2017**, *22*, 65–80. [CrossRef]
76. Yang, L.; Cormican, K.; Yu, M. An ontology model for systems engineering derived from iso/iec/ieee 15288: 2015: Systems and software engineering-system life cycle processes. *World Acad. Sci. Eng. Technol. Int. J. Comput. Electr. Autom. Control Inf. Eng.* **2016**, *11*, 1–7.
77. Drozd, O. Privacy pattern catalogue: A tool for integrating privacy principles of ISO/IEC 29100 into the software development process. In *IFIP International Summer School on Privacy and Identity Management*; Springer: Berlin/Heidelberg, Germany, 2015; Volume 476, pp. 129–140.
78. Culot, G.; Nassimbeni, G.; Podrecca, M.; Sartor, M. The ISO/IEC 27001 information security management standard: Literature review and theory-based research agenda. *TQM J.* **2021**, *33*, 76–105. [CrossRef]
79. Mirtsch, M.; Kinne, J.; Blind, K. Exploring the adoption of the international information security management system standard iso/iec 27001: A web mining-based analysis. *IEEE Trans. Eng. Manag.* **2020**, *68*, 87–100. [CrossRef]
80. Shastri, S.; Banakar, V.; Wasserman, M.; Kumar, A.; Chidambaram, V. Understanding and benchmarking the impact of GDPR on database systems. *arXiv* **2019**, arXiv:1910.00728. [CrossRef]
81. Abouelmehdi, K.; Beni-Hessane, A.; Khaloufi, H. Big healthcare data: Preserving security and privacy. *J. Big Data* **2018**, *5*, 1–18. [CrossRef]
82. Samaraweera, G.D.; Chang, J.M. Security and privacy implications on database systems in Big Data era: A survey. *IEEE Trans. Knowl. Data Eng.* **2019**, *33*, 239–258. [CrossRef]
83. Shmueli, E.; Vaisenberg, R.; Elovici, Y.; Glezer, C. Database encryption: An overview of contemporary challenges and design considerations. *ACM SIGMOD Rec.* **2010**, *38*, 29–34. [CrossRef]
84. IBM. IBM Business Automation Workflow—SQL Server Database Privileges. Available online: https://www.ibm.com/docs/en/baw/19.x?topic=privileges-sql-server-database (accessed on 15 October 2022).
85. Microsoft. SQL Server—Permissions (Database Engine). Available online: https://learn.microsoft.com/en-us/sql/relational-databases/security/permissions-database-engine?view=sql-server-ver16 (accessed on 15 October 2022).

86. Neves, A.L.; Freise, L.; Laranjo, L.; Carter, A.W.; Darzi, A.; Mayer, E. Impact of providing patients access to electronic health records on quality and safety of care: A systematic review and meta-analysis. *BMJ Qual. Saf.* **2020**, *29*, 1019–1032. [CrossRef]
87. Pika, A.; Wynn, M.T.; Budiono, S.; Ter Hofstede, A.H.; van der Aalst, W.M.; Reijers, H.A. Privacy-preserving process mining in healthcare. *Int. J. Environ. Res. Public Health* **2020**, *17*, 1612. [CrossRef]
88. Huang, Q.; Li, H. An efficient public-key searchable encryption scheme secure against inside keyword guessing attacks. *Inf. Sci.* **2017**, *403*, 1–14. [CrossRef]
89. Qian, H.; Li, J.; Zhang, Y.; Han, J. Privacy-preserving personal health record using multi-authority attribute-based encryption with revocation. *Int. J. Inf. Secur.* **2014**, *14*, 487–497. [CrossRef]
90. Samydurai, A.; Revathi, K.; Prema, P.; Arulmozhiarasi, D.; Jency, J.; Hemapriya, S. Secured Health Care Information exchange on cloud using attribute based encryption. In Proceedings of the 2015 3rd International Conference on Signal Processing, Communication and Networking (ICSCN), Chennai, India, 26–28 March 2015; pp. 1–5.
91. Barrett, C. Are the EU GDPR and the California CCPA becoming the de facto global standards for data privacy and protection? *Scitech Lawyer* **2019**, *15*, 24–29.
92. The Department of Health and Aged Care. Electronic Health Records. Available online: https://www.health.gov.au/topics/health-technologies-and-digital-health/about/electronic-health-records#more-about-my-health-record (accessed on 10 January 2023).
93. Madden, C.; Lydon, S.; Curran, C.; Murphy, A.W.; O'Connor, P. Potential value of patient record review to assess and improve patient safety in general practice: A systematic review. *Eur. J. Gen. Pract.* **2018**, *24*, 192–201. [CrossRef]
94. Roehrs, A.; Da Costa, C.A.; da Rosa Righi, R.; De Oliveira, K.S.F. Personal health records: A systematic literature review. *J. Med Internet Res.* **2017**, *19*, e13. [CrossRef]
95. Adebisi, O.; Oladosu, D.; Busari, O.; Oyewola, Y. Design and implementation of hospital management system. *Int. J. Eng. Innov. Technol.* **2015**, *5*, 31–34.
96. Salleh, D.A. Information Systems in Health Care. Available online: https://drdollah.com/hospital-information-system-his/ (accessed on 17 February 2022).
97. Australian Government Office of Parliamentary Counsel. *Healthcare Identifiers Act 2010*; ACT: Canberra, Australia, 2021.
98. Australian Government Office of the Australian Information Commissioner. Healthcare Identifiers. Available online: https://www.oaic.gov.au/privacy/privacy-legislation/related-legislation/healthcare-identifiers (accessed on 12 February 2023).
99. Office of Parliamentary Counsel. *Healthcare Identifiers Regulations 2010*; ACT: Canberra, Australia, 2017.
100. Sajid, A.; Abbas, H. Data privacy in cloud-assisted healthcare systems: State of the art and future challenges. *J. Med. Syst.* **2016**, *40*, 1–16. [CrossRef]
101. Sun, J.; Yao, X.; Wang, S.; Wu, Y. Blockchain-based secure storage and access scheme for electronic medical records in IPFS. *IEEE Access* **2020**, *8*, 59389–59401. [CrossRef]
102. De Angelis, S. Assessing security and performances of consensus algorithms for permissioned blockchains. *arXiv* **2018**, arXiv:1805.03490.
103. Lin, Y.; Zhang, C. A Method for Protecting Private Data in IPFS. In Proceedings of the 2021 IEEE 24th International Conference on Computer Supported Cooperative Work in Design (CSCWD), Dalian, China, 5–7 May 2021; pp. 404–409.
104. Abdullah Lajam, O.; Ahmed Helmy, T. Performance Evaluation of IPFS in Private Networks. In Proceedings of the 2021 4th International Conference on Data Storage and Data Engineering, Barcelona, Spain, 18–20 February 2021; pp. 77–84.
105. Alharby, M.; Van Moorsel, A. Blockchain-based smart contracts: A systematic mapping study. *arXiv* **2017**, arXiv:1710.06372.
106. Mohanta, B.K.; Panda, S.S.; Jena, D. An overview of smart contract and use cases in blockchain technology. In Proceedings of the 2018 9th International Conference On Computing, Communication and Networking Technologies (ICCCNT), Bengaluru, India, 10–12 July 2018; pp. 1–4.
107. Khan, S.N.; Loukil, F.; Ghedira-Guegan, C.; Benkhelifa, E.; Bani-Hani, A. Blockchain smart contracts: Applications, challenges, and future trends. *Peer-to-Peer Netw. Appl.* **2021**, *14*, 2901–2925. [CrossRef] [PubMed]
108. Kang, P.; Yang, W.; Zheng, J. Blockchain Private File Storage-Sharing Method Based on IPFS. *Sensors* **2022**, *22*, 5100. [CrossRef] [PubMed]
109. First Click Testing. Available online: https://www.usability.gov/how-to-and-tools/methods/first-click-testing.html (accessed on 12 February 2023).
110. Design Confidently. Available online: https://usabilityhub.com/ (accessed on 12 February 2023).
111. Altulaihan, E.A.; Alismail, A.; Frikha, M. A Survey on Web Application Penetration Testing. *Electronics* **2023**, *12*, 1229. [CrossRef]
112. Zhou, J.; Feng, Y.; Wang, Z.; Guo, D. Using secure multi-party computation to protect privacy on a permissioned blockchain. *Sensors* **2021**, *21*, 1540. [CrossRef]