cryptography

MDPI

Article

Privacy-Enhancing Technologies in Collaborative Healthcare Analysis

Manar Alnasser [1,2] and Shancang Li [1,*]

[1] School of Computer Science and Informatics, Cardiff University, Cardiff CF10 3AT, UK
[2] Self-Development Department, Deanship of Common First Year, King Saud University, Riyadh 11362, Saudi Arabia
* Correspondence: shancang.li@ieee.org

Abstract: Healthcare data is often fragmented across different institutions (hospitals, clinics, research centers), creating data silos. Privacy-enhancing technologies (PETs) play a fundamental role in collaborative healthcare analysis, enabling healthcare providers to improve care while protecting patient privacy. By providing a compliant framework for data sharing and research, PETs facilitate collaboration while adhering to stringent regulations like HIPAA and GDPR. This work conducts a comprehensive survey to investigate PETs in healthcare industry. It investigates the privacy requirements and challenges specific to healthcare, and the key enabling PETs are explored. A review of recent research trends that identify challenges, and AI related concerns is presented.

Keywords: privacy; privacy enhancing technologies; healthcare; collaboration analysis

Received: 17 February 2025
Revised: 1 April 2025
Accepted: 15 April 2025
Published: 22 April 2025

Citation: Alnasser, M.; Li, S. Privacy-Enhancing Technologies in Collaborative Healthcare Analysis. *Cryptography* **2025**, *9*, 24. https://doi.org/10.3390/cryptography9020024

1. Introduction

Digitisation of health and patient data is significantly changing the clinical, operating, and business models which continue to impact the world of economy for the foreseeable future [1]. In parallel, there is a significant increase in the volume of collected, stored, and processed data. According to The US International Data Corporation (IDC) prediction, that by 2025 the amount of data created will rise to around 163 zetta bytes (ZB) worldwide [2]. Consequently, the vast number of connected devices and the increasing quantity of data fuelling an ever-growing number of applications, are collectively raising significant concerns in related security and privacy issues surrounding this data. While this change has improved patient care workflow and reduced costs, it also increases the probability of security and privacy breaches. In addition, one of the key barriers to widely adopting clinically-validated artificial intelligence (AI) applications is preserving patients' privacy [3].

Based on Statista [4], the healthcare industry is considered one of the most vulnerable to cybercrime. In 2023, it remained the most targeted by cyber-attacks in the US, resulting in data compromises. The number of data compromise incidents in the US increased more than twice compared to 2022. In addition, between 2016 and 2022 the highest number of reported breached records was registered in 2022, totalling 51.4 million. These findings reveal a critical need for healthcare providers to adopt more proactive and comprehensive cyber security strategy to protect against growing cyber threats. In light of these issues, several techniques have emerged to mitigate privacy threats, known as PETs. These technologies that are designed to achieve data sharing and privacy preservation are gaining an expanding interest.

In the past decade, there is a significant interest in using PETs technologies to enhance privacy in healthcare industry. A comprehensive literature review was conducted to pro-

vide researchers with an overview of PETs in healthcare and identify research opportunities. It is based on data from three major academic databases (*SCOPUS, IEEE, and ScienceDirect*) illustrating the increasing adoption of PETs in healthcare environment. Figure 1 indicates the number of recent publications in PETs in healthcare industry. The numbers include studies in PETs in general, as well as specific technologies including *Data Minimisation, Anonymization, Pseudonymization, Homomorphic Encryption (HE), Secure Multiparty Computation (SMPC), Differential Privacy, Zero-Knowledge Proofs (ZKP), and Federated Learning*. As a result, this paper investigates PETs and then provides a brief overview of key enable technologies in healthcare field, including summary of their challenges, to encourage future research to enhance privacy.

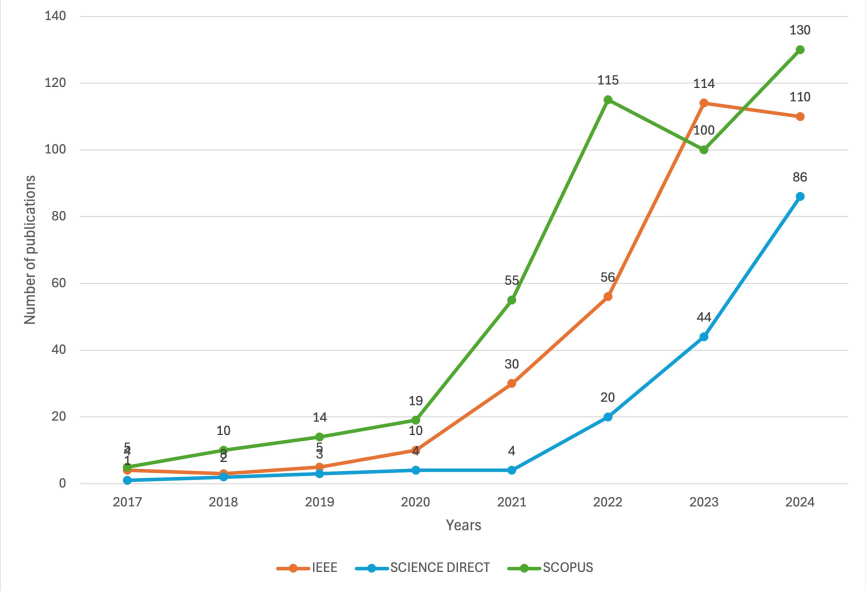

Figure 1. The number of publications in PETs in healthcare industry.

Aiming at providing a comprehensive overview of the use of PETs in the healthcare industry, this work reviewed the most recent PET solutions. This survey addresses three guiding research questions: (1) Which key PETs are currently deployed in a healthcare system, and how do they function? (2) What privacy requirements and challenges define their application in this domain? (3) What barriers hinder their widespread adoption, particularly regarding data utility in collaborative analysis? The main contributions can be summarised as

(1) A comprehensive literature review has been conducted to focus on recent research studies using PETs in healthcare systems and investigations of the privacy requirements and challenges in healthcare industry.
(2) This work investigates key enabling PETs, including federated learning, differential privacy, homomorphic encryption, synthetic data generation, multi-party computation (MPC), etc., and analysed how they affect the data utility in collaborative analysis.
(3) Key recent research trends in the protection of healthcare data analysis were addressed, specifically highlighting privacy protection schemes within AI models utilizing healthcare data, and their impact on data utility.

2. Related Works

Numerous comprehensive surveys have examined various aspects of PETs in different fields such as AI and IoT. Collectively, these surveys offer a detailed understanding of PETs

evolution and current trends. By summarising key insights from these reviews, existing gaps are identified, that provide a basis for the contributions presented in this paper.

Cha et al. conducted a comprehensive survey on PETs in internet of things (IoT) applications via a newly proposed categorization [5]. In addition, the works in [2,6] focused on classifing PETs into classes, whereas [7] proposed a framework for PETs application in IoT communications and evaluated according to stakeholder and GDPR requirements. In smilier context, a paper investigated the privacy issues of individual privacy in the healthcare industry related to wearable IoT and reviewed the international guidelines and laws that support it [8,9]. This work is an endeavor to investigate the use of PETs to enhance privacy in collaborative healthcare analysis. Recently, PETs technologies have been increasingly employed to enhance the privacy in various environments, and a number of new solutions have developed specifically to improve the privacy in collaborative healthcare analysis. Due to the healthcare ecosystems consisting of many dimensions, vulnerability could be exploited to undertake several attacks and breach privacy. To ensure a comprehensive review, this paper systematically surveyed literature from sciencedirect, IEEE Xplore, and Google scholar, using keywords such as 'privacy enhancing technologies' and 'PETs in healthcare'. It focused on peer-reviewed studies from 2018–2024, selecting papers that applied PETs and provided empirical insights. Table 1 summarise related works, including their focus area, contributions, and limitations.

Table 1. Summary of existing works of PET methods.

Ref.	Methods	Strengths	Limitations
[6]	Various PETs	the legal foundation of PETs and provided a classification of PETs and a selection of some of the most relevant PETs.	Economic, social and usability aspects of PETs.
[10]	Various PETs/IoT area	Assess the development of PETs across fields, evaluating their compliance with legal standards and effectiveness in mitigating privacy threats.	Need research of PETs in the category of holistic privacy preservation
[11]	Various PETs/IoT area	Analyse, evaluate, and compare various PETs that can be deployed at different layers of a layered IoT architecture to meet the privacy requirements of the individuals interacting with the IoT systems.	A careful consideration of the unique features associated with the IoT, including the use of heterogeneous power-limited devices and the massive need for streaming data flow
[2]	Various PETs	A taxonomy classifying eight categories of PETs into three groups, and for better clarity.	Point out which PETs best fit each personalized service category. The trade-off between privacy preservation and personalized services, Technical, user experience, legal, and economic challenges.
[7]	Various PETs/IoT area	A framework for the application of PETs in IoT communications. discuss an example implementation based on a car-sharing service.	Develop a security model for the framework. Possible threats include, e.g., rogue framework instances and malicious traffic injection.
[12]	Various PETs/Blockchain	present PETchain: a novel privacy enhancing technology using blockchain and smart contract.	Checking PETchain compatibility with GDPR to improve it.

Table 1. *Cont.*

Ref.	Methods	Strengths	Limitations
[13]	Various PETs	Investigates several industrial use cases, their characteristics, and the potential applicability of PETs to these.	Handle large volumes of data and address requirements.
[11]	Federated Learning/Healthcare	Take Alzheimer's disease (AD) as an example and design a convenient and privacy-preserving system named ADDETECTOR with the assistance of Internet of Things (IoT) devices and security mechanisms.	Discover more effective features to represent the characteristics of ADs and evaluate the feasibility of ADDETECTOR on a larger dataset.
[14]	Various PETs/IoT area	Reveal the landscape of PETs in data markets for the IoT. Identify and filter the studies aiming to solve this landscape's challenges.	The IoT challenges for privacy enhancement, consequences of a lack of interoperability, computation and storage constraints, and the privacy disparity across jurisdictions.
[15]	SMPC, HE, DP, CC	a detailed analysis of collaborative ML approaches from a privacy perspective, and a detailed threat model and security and privacy considerations for each collaborative method. Deeply analyse (PETs) in the context of collaborative ML.	Verifiability of computations to provide proof points in collaborative ML/AI message flow
[16]	FL/Smart Healthcare	Review on the emerging applications of FL in key healthcare domains, including health data management, remote health monitoring, medical imaging, and COVID-19 detection. Analyse Several recent FL-based smart healthcare projects	Communication Issues in FL-based Smart Healthcare. Standard Specifications for Federated Healthcare Deployment. Security Issues in FL-based Smart Healthcare
[17]	Various PETs/healthcare	An overview of how to integrate PETs into pandemic preparedness	Privacy/Utility Trade-Off, Infrastructure Deployment, Public Trust and Acceptance.
[18]	DP/Smart Home Data	Employ the Local Differential Privacy (LDP) technique and propose a framework for securing data collection in smart homes based on the k-Anonymity Ran-domized Response (k-RR) algorithm.	Explored alternative probabilistic models, such as the Maximum Entropy Markov Model (MEMM), Gaussian distribution, or Dirichlet distribution, for comparative purposes.
[19]	DP/IoMT	A Group-based DP (GDP)framework for Process Mining to protect the privacy of healthcare data in specific columns which are neither activity nor class ID. evaluate of prominent PM algorithms.	Striking a balance between DP and data utility in PM poses a challenge. Resource optimization
[20]	FL, AI/IoT	Investigate developments in FL for edge AI, with an emphasis on strengthening security and resilience against adversarial attacks.Examine privacy-preserving methods. Explore personalization techniques that enable FL models to adjust to the unique needs of individual IoT devices, enhancing system performance and user experience.	Data Heterogeneity, Communication Efficiency, Privacy and Security, Scalability and Resource Constraints, Personalization and Model Adaptation, AND Incentive Mechanisms

Since FL is designed to solve the problem of privacy data protection in machine learning, it faces challenges in ensuring privacy protection. It is important to ensure that the training model does not expose users' private information in FL process. Another challenge

is the insufficient amount of data. A high performance model training in traditional machine learning required a large amount of data in a distributed environment, but the amount of data on each mobile device is insufficient [21]. Another complex challenge is systems heterogeneity. FL must support multiple devices with different hardware and configuration, resources, and operating system to achieve robust distributed learning performance [22]. Also, Although FL has decentralized nature and able to address some privacy concerns by maintaining data on local devices, but there is a security risks such as data poisoning, model inversion, and adversarial attacks. FL has been adopted in initiatives like MELLODDY which was comprised of 10 pharmaceutical companies, academic research labs, large industrial companies and startups apply FL to drug discovery by creation of a global federated model for drug discovery without sharing the confidential data sets of the individual partners [23]. Yet, FL faces model inversion attacks, where attackers can reconstruct patient features, posing risks in small-scale healthcare settings. Defences like DP in healthcare applications, minimize this but reduce utility.

Privacy-Enhancing Technologies

In the past decade, the field of PETs has evolved rapidly with diverse approaches. Figure 2 summarises the key taxonomy of key PETs based on their usage. The choice of technique is determined by resource availability and the nature of the privacy issues. However, there are limitations to the adoption of PETs in general applications due to many obstacles varying from degradation of utility of data, to high costs of computational and communication [24].

The initial introduction of the term "Privacy-enhancing Technologies" was in 1995 in a report on PETs, which was published by Dutch Data Protection Authority in collaboration with the Privacy Commissioner in Ontario/Canada that investigated a novel approach to privacy protection [6].

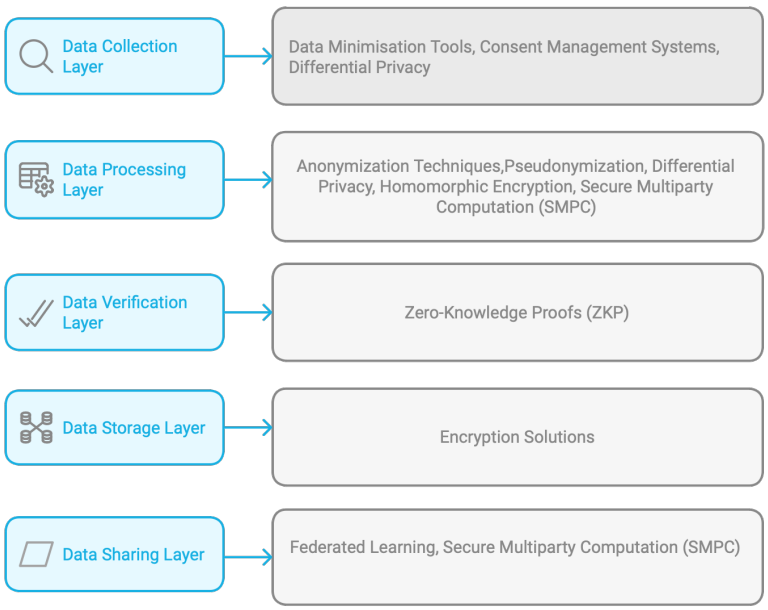

Figure 2. The taxonomy of key PETs based on their usage.

There is a growing body of literature that recognizes the importance of PETs in different fields. In 2017, Fischer-Hbner & Berthold [6] defined the legal foundation of PETs and classified PETs into three classes based on the legal privacy principles. In 2020, Kaaniche et al. [2] classified PETs into three different groups and eight categories, based on the main entity involved in the privacy-preserving decision. The groups are user-side techniques,

server-side techniques, and channel-side techniques. They also identified the potential of PETs to satisfy both usually divergent economic and ethical purposes. Cha et al. conducted a comprehensive literature review on PETs in IoT applications using a newly proposed categorization [5].

It outlined the current status, limitations, and future research directions related to PETs in the IoT. It found that existing PETs in the IoT are not advanced enough to fully align with the principles outlined by the GDPR and the ISO 29100 standard [25]. On average, each research paper addressed six privacy principles. Kunz et al. have presented a framework for PETs application in IoT communications. It identified stakeholder and GDPR requirements and evaluated the framework design based on these requirements to demonstrate its ability to support requirements such as data minimization and data protection by design. In addition, an example implementation based on a car sharing service was presented and discussed [7]. A study published in 2022 by Li et al. develop a privacy-preserving smart healthcare system for low-cost Alzheimer's disease (AD) detection. In the system, the audio from smart devices are used as the input and the differential privacy based mechanism and federated learning based framework are applied to prevent the leakage of raw data and model details during transmission. The experimental results proved that AD Detector achieves a high level of accuracy while ensuring strong security protection [26].

In 2020, Terhorst et al. proposed privacy evaluation protocols (PEPs) for the evaluation of Soft-Biometric PETs. The framework assessed PETs in the most critical attack scenario, where a function creep attacker that knows and adapts to the systems privacy-mechanism. They proposed three PEPs to ensure that the data is used appropriately, considering the nature of the PETs evaluated. It is based on Kerckhoffs's principle of cryptography to guarantee that the protocol meets the highest standards in both cases [27]. In the same context, Haddad proposed a privacy-preserving handover protocol that utilizes blockchain and zero knowledge poof (ZKP) to enhance privacy and security in 5G networks [28] . It ensures a seamless and secure transition for user equipment, maintains confidentiality, and strengthens the defense of the network against potential adversaries. Gatha et al. investigated the existing PETs and their applicability to real-world data. It established that PETs depend on the nature of the data and its intended use. It explored the combining both syntactic and semantic PETs applicability and shown the effectively of comprehensive approach [29]. In 2021, Javed et al. proposed a PET using blockchain and smart contract [12]. The technique aims to comprehensively address user privacy by allowing users to define their access control policy by deploying their smart contract. Garrido et al. studied PETs challenges and provided guidelines synthesized from expert interviews and a literature review focus mainly on automotive use cases in 2021 [13].

A study conducted by Soykan et al. in 2022 noted that deeply analyzed PETs, including secure multiparty computation, homomorphic encryption, differential privacy, and confidential computing in the context of collaborative machine learning [15]. It presented a guideline for selecting privacy-enhancing technologies for collaborative machine learning and privacy practitioners. This study is the first survey aimed at providing an in-depth focus on collaborative ML requirements and challenges for privacy solutions while also offering guidelines for PETs selection. Then, an extensive literature review of PETs applicable for collaborative machine learning is provided. PETs establish the foundation for pandemic preparedness by addressing social concerns related to data collection and analysis, thereby supporting various data analysis and learning tasks for future application scenarios. Liu et al. presented an overview of how PETs can be integrated into pandemic preparedness and proposed a vision of future directions for leveraging more data for pandemic response during public health emergencies, ensuring privacy, effectiveness, efficiency, and explainability [30].

In 2023, Waheed et al. proposed a framework for securing data collection in smart homes based on the k anonymity randomized response (k-RR) algorithm to ensure comprehensive security for data generated by smart homes. This approach achieved a dual layer of privacy protection, addressing the security concerns associated with IoT devices in smart cities. The results demonstrated the effectiveness of the proposed framework in evaluating the privacy risk of both obfuscated and original high-risk home data [18]. Sahlabadi et al. presented a Group-Based DP (GDP) framework for process mining to protect the individuals privacy, while enabling healthcare organizations to extract valuable insights from their logs [19]. Miller et al. paper explored data privacy research as applicable to database programming curriculum in information technology. It outlined the differential privacy concepts and definitions, used to obfuscate sensitive data collected and stored in modern relational database schemes. It provided knowledge for integrating the differential privacy concept and closely related PETs into undergraduate level relational database management system class with focus on traditional business model [31].

This highlights the critical need to identify privacy requirements and challenges for wearable IoT (WIoT) health devices, particularly given the vast amounts of sensitive user data they collect. Li et al. comprehensively reviewed the state-of-the-art, considering IoT architecture, current privacy laws, and representative PETs. They analyzed, evaluated, and compared various PETs applicable at different layers of the IoT architecture to meet individual privacy needs. Their findings emphasized the importance of carefully considering the unique characteristics of IoT when adopting existing PETs, while acknowledging the significant potential of many PETs for IoT applications [32]. Nguyen et al. provided a detailed review of FL usage in smart healthcare and discussed recently advanced FL designs and the key applications that would be useful to federated smart healthcare. Moreover, it highlighted health data management, remote health monitoring, medical imaging, and COVID-19 detection [16]. In the same context, in 2024, Rane et al. investigated federated learning for edge AI development, focusing on enhancing security and resilience against adversarial attacks like model inversion and data poisoning. It examined privacy-preserving methods such as homomorphic encryption and differential privacy to ensure the safety of private information [20].

3. Privacy Requirements and Challenges in Healthcare Industry

Modern healthcare systems rely heavily on advanced technologies like IoT devices and cloud computing to collect and analyse personal health data at an unprecedented scale. While these analytics offer significant benefits, such as remote patient monitoring, early disease diagnosis, and personalized treatments, they also raise serious privacy concerns. Without robust safeguards, this data analysis can become a privacy nightmare. The protection of privacy is as essential as the development and delivery of quality healthcare services. The promised services will not be delivered as expected without privacy. This section will highlight privacy challenges and requirements in the healthcare industry.

In [8], the authors identified key challenges and requirements for wearable IoT (WIoT) health devices, outlining four essential capabilities for IoT applications. These include: (1) seamless and secure connectivity with other devices for managing device functions and encrypting data; (2) efficient power management for long-term, uninterrupted monitoring, a critical design consideration; and (3) user-centric design, prioritizing comfort and ease of wear. To meet these requirements, the devices should be small and lightweight. The last challenge is that the collected data should be stored securely in the microcontroller with large memory to avoid data loss in case of disconnection. Moreover, it summarized the privacy requirements for WIoT in the healthcare industry based on many studies and researches as follows: data minimization, user participation, data anonymization,

authentication and authorization, level of availability, efficiency, and effectiveness of privacy capabilities, limit the personally identifiable information (PII) processing, the right to process a patient's own protected health information (PHI) by the patient only anytime, users right to limit usage of their health data, the right to complain, and Hide and encrypt the personal data from plain view. Hari et al. study highlighted integrity and scalability as challenges related to data privacy [33].

Tandon et al. explained and conducted a comparative study of security and privacy challenges in healthcare identified key requirements, including confidentiality, integrity, authentication, availability, data freshness, data maintenance, non-repudiation, interoperability, privacy, fault tolerance, and secure booting. It also highlighted several security and privacy challenges including data management, shield architecture, authentication, computational and memory limitations, authorization and access control, eavesdropping, data leakage and destruction, trust mechanism, secure network, and scalability [34].

Moreover, it classified IoT systems to three-layer as perception, transportation, and application, and identified issues and weaknesses in the security at each layer. The perception layer is vulnerable to physical attacks, impersonation, denial-of-service, routing attacks, and data transit attacks. The transport layer is susceptible to routing attacks, denial-of-service, and data transit attacks. The application layer is vulnerable to data leakage, denial-of-service, and malicious code injection [35]. Louassef et al. identified two key requirements for maximizing patient data privacy in IoT systems: content requirements and contextual requirements [36].

- **Content Privacy**: ensures and preserves patient data to prevent attackers from revealing it. However, this is insufficient for robust privacy, as attackers can potentially identify patient data by targeting the receiving doctor's identification.
- **Contextual Privacy**: this involves two distinct sub-requirements: pseudonymity, where pseudonyms are used in lieu of real identities; and anonymity, which goes further by ensuring that patient identities remain unidentifiable from their data or actions. Anonymity includes preserving both patient and medical anonymity, along with unlinkability and unobservability.

Zhang et al. summarized the privacy requirements for healthcare blockchains, encompassing user anonymity, patient control over data, confidential transactions, fine-grained access control to transactions, and user authentication [37]. In 2021, a study conducted to investigated requirements for security, privacy, and trust of the precision health and categorised them into three requirements as follow law requirements, ethics requirements, and health domain requirements. The study identified a number of major requirements from regulatory aspects, including a proper informed consent, secure and privacy of data processing, data transfer security, proper administrative, confidentiality, physical and technical safeguards, integrity checks, transparency, availability, fairness, breach notification, and minimal and limited data use. From ethical perspectives, key considerations include obtaining ethics approval, raising awareness, granting control to individuals, maintaining ownership, minimising information leakage, managing data-sharing properly, and conducting risk analysis and management. These ethical requirements along with regulations requirements, including privacy, security, confidentiality, trust, and breach notification, represent the essential requirements.

Moreover, while the sensitivity of PH data usage in health decision making has critical role, the trustworthiness of health data is a particular requirement. This requires delivering precision health data in a standard format that is simple, clear, complete, accurate, timely, and transparent to guarantee the effectiveness and correctness when making inferences for decisions. Based on these requirements, it identified four significant challenges; they are

health data security and privacy whilst computing, consent management, precision health data trustworthiness, and legal and ethical compliance [38].

Figure 3 illustrate the essential requirements and challenges associated with privacy in the healthcare industry, classifying requirements into regulatory, ethical, and technical requirements and essential challenges.

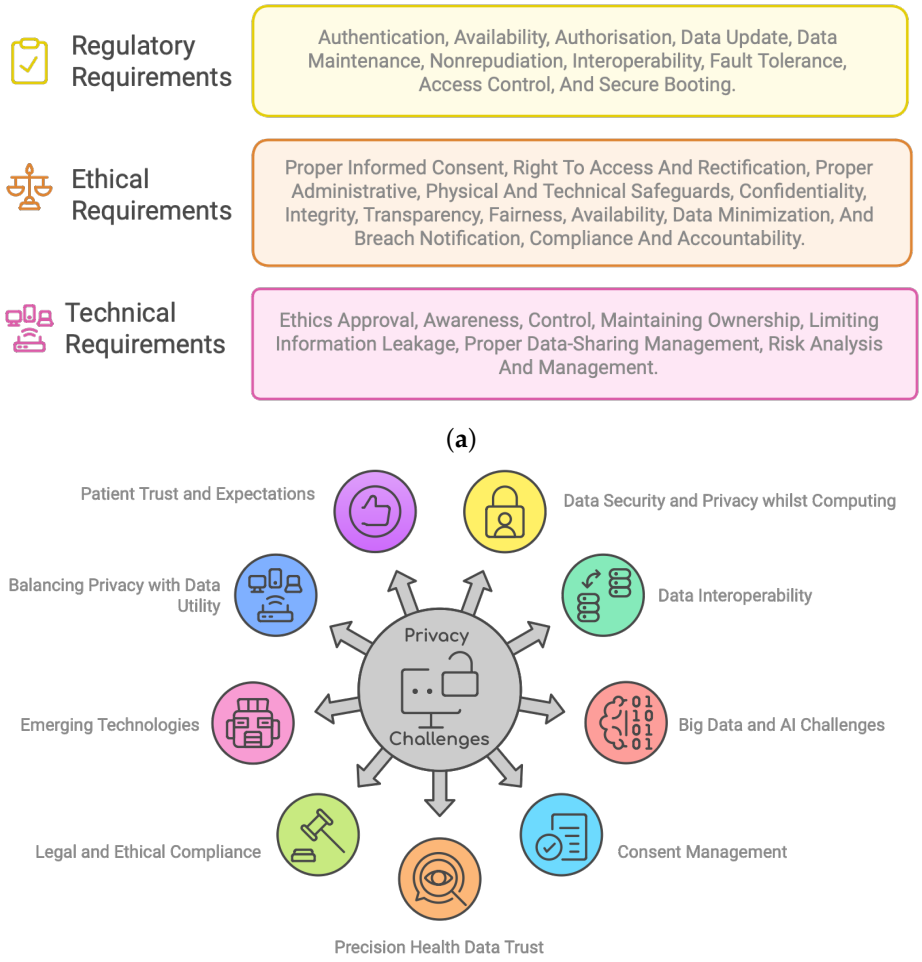

Figure 3. Privacy Requirements and Challenges in Healthcare industry. (**a**) Essential Privacy requirements; (**b**) Key challenges.

4. Key Enabling Privacy Enhancing Technologies (PETs)

This section provides an overview of key PETs that can be used to mitigate privacy attacks. It focuses on the following PETs: *data minimisation, federated Learning, homomorphic encryption, and anonymization*.

4.1. Data Minimisation

Data minimization, the practice of collecting and using only the personal data necessary for a specific purpose, is a core principle of many privacy regulations, such as GDPR, HIPAA, and CPRA [39,40]. These regulations dictate that only required data be collected to fulfil a certain purpose [41,42]. In the literature, data minimisation is commonly defined with phrase such as: specifying the purpose of data processing when the data is collected, deleting data when no more required for the specified purpose, limiting the amount of shared data to the minimum required, and to minimise collection of personal data. similar phrases found in privacy regulatory documents, For example, HIPAA states “to limit the

scope of the PHI (Protected Health Information) the health systems use, disclose or request to the minimum necessary".

Data minimization aims to limit personal data collection and retention to the minimum necessary, ensuring its use and sharing are purpose-specific. However, existing definitions often reiterate regulatory language, providing little practical guidance on implementation [43].

Data minimization poses a challenge for system designers. Developers must ensure data usage is adequate, relevant, and limited to the system's specific purpose. However, modern software systems often rely on vast datasets, not all of which may be directly related to the initial collection purpose. Consequently, developers frequently find data minimization the most challenging privacy regulation to implement, often citing conflicts with business interests [44]. A recent study revealed that while software developers often prioritize minimizing data storage and sharing when implementing data minimization, they struggle to effectively minimize data use within these processes. This challenge arises from the complex and sometimes counterintuitive nature of implementing data minimization in system design. That is, developers lack a clear foundation logic to guide their decisions to avoid the use of a particular data item [44].

4.2. Federated Learning (FL)

A federated learning has emerged as a powerful machine learning approach aims to protecting the privacy of data. It based on a principle of training machine learning models on decentralized entities holding local data without sharing them [45,46]. It was introduced by Google in 2016 for updating models for Android mobile terminal users locally [47]. Since then, FL has received significant attention both in academic and industrial fields [48]. Instead of sharing the raw data to a centralized location, only the local models are updated and exchanged between a parameter server and the clients. This decentralized approach ensures the security of sensitive data, minimises the risks of unauthorized access, or data breaches. Consequently, each entity can benefit from other entities, enhancing the accuracy of model.

In recent years, there have been variety applications adopted FL in practice, such as prediction loan status, health situation assessment, and next-word prediction [49,50]. FL should include multiple users collaborate to build a shared machine learning model. Each user participates in training by using their local data. During the process, the data will not leave the local client. The relevant information of model such as model structure, and model parameters can be securely shared in an encrypted format. Moreover, the performance of FL model should be able to approach the ideal model performance, that is, the model that collects all training data and trains it [51]. FL can be classified based on distribution of data as horizontal federated learning, vertical federated learning, and federated transfer learning [52,53].

Horizontal FL is appropriate when datasets share a substantial number of user features but have limited user overlap. Vertical FL, on the other hand, is designed for situations where datasets share few features but have significant user overlap. When both user and feature overlap is minimal, transfer learning can be used to address data or label scarcity [21]. Although a relatively new technology, FL has experienced rapid adoption due to the increasing prevalence of deep learning applications. Healthcare, where patient privacy is crucial, is a prime example of an industry benefiting from FL [9,54].

FL has recently been described as a key factor for the digital health future [49]. It supports collaborative data sharing and analytics in health research by enabling rapid responses to emerging threats and facilitating real-time international data exchange and analysis [55]. Since FL is designed to solve the problem of privacy data protection in machine learning, it faces challenges in ensuring privacy protection. It is important to

ensure that the training model does not expose users' private information in FL process. Another challenge is the insufficient amount of data. Training of a high performance model in traditional machine learning required a large amount of data in a distributed environment, but the amount of data on each mobile device is insufficient [21]. Another complex challenge is Systems heterogeneity. FL must support multiple devices with different hardware and configuration, resources, and operating system to achieve robust distributed learning performance [22].

4.3. Homomorphic Encryption (HE)

HE allows computations on encrypted data without decryption [56]. Using public-key cryptography, data is encrypted with a public key, and only the corresponding private key can decrypt it. HE ensures that identical mathematical operations on encrypted and decrypted data yield equivalent results [56]. HE schemes vary in their supported operations. FHE handles both addition and multiplication, while PHE supports only one, offering significantly better performance. Schemes supporting a limited set of operations beyond PHE are classified as somewhat homomorphic encryption (SHE) [14,17].

HE involves four key steps: (1) Key generation, producing public and private keys for encryption and decryption; (2) Encryption, converting plaintext to ciphertext using the public key; (3) Decryption, recovering the plaintext from ciphertext using the private key; and (4) Evaluation, performing operations on ciphertext while maintaining its format [17]. The major limitation of Fully Homomorphic Encryption is its computation overhead, due to its high computational complexity and the comparatively large storage requirements for its cyphertext. This performance issue is considered a significant challenge leading the research projects to adopt PHE instead of FHE [9]. The major limitation of Fully Homomorphic Encryption is its computation overhead, due to its high computational complexity and the comparatively large storage requirements for its cyphertext. This performance issue is considered a significant challenge leading the research projects to adopt PHE instead of FHE [9]. HE is computationally intensive, particularly when applied to large datasets or complex computations, which lead to significant processing overhead and slow in application performance. This can be a critical barrier in time-sensitive scenarios, such as real-time data analysis or clinical decision-making. Additionally, a deep understanding of advanced mathematical concepts and algorithms are required to implement it. Moreover, It has limited functionality, and does not support all types of computations effectively. Operations, such as non-linear computations, are challenging to perform which restrict the applications range that can leverage HE effectively.

Although HE is very slow in performance and complex [8], it has the ability to preserve sensitive information while maintaining a level of service [57,58]. It is one of the most advanced PETs but still not capable of covering all real-world use cases [9]. AS A real-world implementations of HE, IBM and Cleveland Clinic used it to enable secure AI analysis of encrypted patient records, researchers can conduct studies without exposing sensitive data. In addition, Biometric identifiers like human DNA and RNA sequences similar to fingerprint and can disclosure a critical information such as disease risk or socially for the identification of family of diseases, such as the presence of an Alzheimer'sallele or the discovery of non-paternity. However , current strategies for genomics data protection are come with high overhead on researchers. Homomorphic encryption can be highly suitable for genomics data sharing [10].

4.4. Anonymization

Data anonymization employs various techniques to protect private or personal information during data collection (e.g., relational, graph-oriented). This involves removing,

altering, or encrypting identifiers that could directly reveal identities or link data to specific individuals or entities [9,14]. There are three commonly accepted data privacy constraints, K-Anonymity L-Diversity, and T-Closeness [1].

Insufficient data anonymization may cause disastrous consequences and increase identity disclosure or re-identification risks [59]. While some PETs protect sensitive data from unauthorised access and ensure confidentiality while maintaining data and computation integrity, the authorized receiver of the plaintext may still be able to reverse engineer the output and link data records to individuals resulting in a re-identification attack. Consequently, applying only secure and outsourced computation PETs is insufficient to achieve the required level of privacy in cases where the receiver may not be fully trusted. Anonymization technologies offer a solution in these situations by preserving implicit identifiers and sensitive attributes [14].

Data anonymisation is considered as the most mature technology developed, with a high degree of usability, and the simplest approach at both theoretical and technical level. It is widely accepted and commonly used in the real world, with sharing anonymised datasets being routine. An ongoing debate questions the sufficiency of data anonymisation techniques in protecting the privacy of entity, with several real-world cases highlighting successful reidentification attacks in healthcare [60]. Nevertheless, anonymisation processes need to be aligned with the nature, scope, context, and processing objectives, while considering the potential risks to the rights and freedoms of natural person, which requires some specialised expertise. Current legislation takes anonymisation into account, significantly reducing the legal and bureaucratic obstacles when handling anonymised data [9].

The anonymization technologies can be categorized into multiple groups; the first one is syntactic technologies which assign a numerical value to the level of individuals' protection in a dataset, leading to a notable perturbation of data, so it is harder to distinguish between the three individuals for an attacker [61]. Semantic technologies is the second category of anonymization technologies, enforce a privacy definition to a learning mechanism used on a dataset, namely differential privacy, ensuring that the output distribution of the mechanism unaffected by the removal or addition of an individual in the dataset. Semantic technologies offer an advantage over syntactic technologies by providing a mathematical guarantee of privacy independent of any background information, so an attacker unable to use related information to re-identify an individual in the dataset. Other anonymization technologies are perturbation and pseudonym creation [14,62]. Table 2 Privacy enhancing technologies summaries.

Table 2. Key PET Technologies.

PET	Description	Use Cases	Strengths	Limitations
Data minimization	Restricting personal data collection and use to the minimum necessary	Privacy-by-design systems, Data compliance	Legal compliance, High level of privacy	Loss of utility, Conflict with business interests, Complex
FL	Training machine learning models on decentralized entities holding local data without sharing them	AI development, Mobile app, Healthcare	Data not shared, Decentralization, Scalability	Insufficient amount of data, Privacy concerns, Systems Heterogeneity
HE	Allows performing computations on encrypted data instead of raw data	Financial analysis, Healthcare, Cloud computing	Data encrypted throughout process, Compatible with most data types	Computation overhead, Complex
Data anonymization	Techniques to protect personal information during data collection	Public dataset research, Healthcare	Cost and risks reduction, Easy to implement	Risk of re-identification, Loss of utility

5. Discussion and Future Works

This Section presents a critical discussion about previous enabling PETs techniques and points to promising research directions that have not yet been widely considered as a future works.

5.1. Discussion

The sheer magnitude of data, technologies, and systems makes achieving privacy in any industry a gargantuan, and at times seemingly impossible task. However, there are guidelines and research directions that can be followed to significantly improve the level of privacy, especially in the future healthcare industry. This does not necessarily require creating new PETs but rather applying existing ones on a large scale, effectively adopting a more holistic approach. This section highlights challenges of PETs adoption in the context of healthcare industry.

This surveys shows the complexity of adopting PETs including a large number of challenges. PETs have been considered as an expensive innovation. The computational cost and overhead play a significant role when organizations consider adopting PETs especially if AI or ML get involved. As AI/ML systems become more advanced, the need to balance privacy with performance and cost-efficiency becomes increasingly important. Implementing a PET platform generally requires extensive software, hardware, and systems integration investments. Moreover, timing of PETs implementation is critical. Integrating PETs during new system deployments is generally cost-effective and often the only practical approach. Retrofitting existing systems with PETs is considerably more complex and costly, exceeding the expense of traditional measures. Consequently, computational cost and overhead remain significant concerns. Furthermore, balancing privacy and data utility presents a major challenge, requiring professionals to carefully weigh individual privacy enhancements against the need to preserve data utility. This challenge is the main cause of the conflict between data owners seeking to maximize privacy and consumers aiming to maximize utility, which is often determined by authenticity of the data. Furthermore, managing the privacy-utility trade-off at each stage of information flow—input, computation, output, transit, and storage—increases complexity for privacy officers. Conversely, the quality of decisions made by data scientists and other decision-makers relies on the integrity of computations and the authenticity of data and identities, factors directly influenced by this same trade-off. PETs are often not adopted because of concerns that they could degrade data quality and negatively impact the quality of the service provided.

Regarding to data in healthcare industry, the analysis of patient data from multiple sources can help in diagnostic and therapeutic decisions but, the exchange of information across multiple parties represent a challenge to apply PETs due to systems and datasets heterogeneous nature. Differences in systems and datasets, arising from their varied sources may affect in PETs performance, indicating a critical need for extensive preprocessing and data homogenisation. Additionally, an important factor that should be considered when using PETs is the need to combine multiple PETs technique altogether instead of adopting only one of them to achieve an acceptable level of privacy. but would also come with its own specificity and additional costs. Clearly, this approach needs a further investigated regarding its applicability, maturity, potential limitations and effective implementation strategies. In addition, security concerns represent one of the major challenges of deploying PETs especially which related to AI. These concerns including data poisoning which occurred when attackers inject malicious data into a training dataset to cause the AI model to produce inaccurate results or degrade its overall performance. Additionally, model inversion attacks are another concern, involve trying to reconstruct private information, such as medical records of individuals, by exploiting the AI model output. attackers also can

create specialised inputs with the purpose of confusing AI model to generate an inaccurate prediction which, a technique referred to as adversarial Attacks. Model extraction represents another concern when the attacker tries to steal an AI model by extract its parameters and leveraging its responses to steal and replicate a model. This poses significant risks, particularly if the model contains confidential or sensitive information [63].

Aside from these challenges that arise during PETs adoption, there are multiple issues encompassing technical, regulatory, and ethical domains, including inherent complexity, low maturity of some PETs, compatibility issues with legacy systems, scalability issues, and regulatory compliance all of which need to be addressed and resolved.

5.2. Future Works

The future work for PETs appears promising, with advancements driving innovation in response to heightened awareness around data privacy. This section highlighted some emerging trends and potential research directions that could further enhance the PETs landscape.

(1) **Secure AI models training.** Develop methods for training AI models in a way that protects the privacy and security of the data, and the models themselves. This approach is important specially in industries that involve sensitive data, like finance, healthcare, and national security. This study will investigate potential vulnerabilities in model updates and ensure the models integrity while preserving privacy during the training process. The aim is to train accurate AI models without compromising the confidentiality of the data or the model parameters.

(2) **Hybrid privacy enhancing technologies.** Construct hybrid solutions that adopt multiple PETs instead of using only one of them. The goal is to enhance the diverse requirements of data privacy in increasingly digital world. Clearly, further investigations are need in this approach in regard to its applicability, maturity, potential limitations and effective implementation strategies. An example of two promising techniques that can be used to address this privacy issue are HE and DP. HE enables secure computations on encrypted data, while DP offers strong privacy guarantees by adding noise to the data.

(3) **Lightweight PETs development.** Developing lightweight versions of HE, SMC, or DP that could be applied especially in real-world AI and machine learning environments. These advancements will allow organizations to leverage the benefits of AI while persevering robust privacy protection. This work would be especially significant in industries that have strict privacy requirements, such as healthcare and finance, where data protection is critical.

(4) **Advances in Cryptographic Techniques.** Advances in cryptographic techniques, especially FHE, are crucial. While FHE offers strong security and privacy, its computational demands are significant. Balancing privacy with computational efficiency is essential for practical real-world applications. Research in cryptography is anticipated to lead to the development of sophisticated methods for data security.

6. Conclusions

As the digital era progresses, many opportunities brought to the healthcare industry, supported by technology and IoT platforms, making e-healthcare more common than ever. However, in the term of data privacy and data collaboration, one of the most sensitive industry is the healthcare industry. The sensitive nature of healthcare records makes them vulnerable to many attacks which in turns make data privacy a growing importance issue in healthcare industry . Therefore, this work presented a comprehensive literature review that focuses on recent research studies using PETs in healthcare systems exploring the

privacy requirements and challenges in healthcare industry. It investigates key enabling PETs. In addition, recent research trends on PETs are summarized and identified challenges, AI related concerns, and data utility in healthcare environments. The result from this study indicated that due to the growing in healthcare industry, it is crucially needed for using PETs techniques in order to enhance the privacy and ensure data integrity, and availability.

Author Contributions: Conceptualization, M.A. and S.L.; writing—original draft preparation, M.A.; writing—review and editing, S.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

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

References

1. Khatir, R.A.; Izadkhah, H.; Razmara, J. Clustering-Based Anonymization Technique using Agglomerative Hierarchical Clustering. In Proceedings of the 2022 8th Iranian Conference on Signal Processing and Intelligent Systems (ICSPIS), Behshahr, Iran, 28–29 December 2022; pp. 1–5.
2. Kaaniche, N.; Laurent, M.; Belguith, S. Privacy enhancing technologies for solving the privacy-personalization paradox: Taxonomy and survey. *J. Netw. Comput. Appl.* **2020**, *171*, 102807. [CrossRef]
3. Khalid, N.; Qayyum, A.; Bilal, M.; Al-Fuqaha, A.; Qadir, J. Privacy-preserving artificial intelligence in healthcare: Techniques and applications. *Comput. Biol. Med.* **2023**, *158*, 106848. [CrossRef]
4. Number of People Affected by Health Data Breaches U.S. 2022. Available online: https://www.statista.com/statistics/798564/number-of-us-residents-affected-by-data-breaches/ (accessed on 9 July 2024).
5. Cha, S.C.; Hsu, T.Y.; Xiang, Y.; Yeh, K.H. Privacy Enhancing Technologies in the Internet of Things: Perspectives and Challenges. *IEEE Internet Things J.* **2019**, *6*, 2159–2187. [CrossRef]
6. Fischer-Hbner, S.; Berthold, S. Privacy-enhancing technologies. In *Computer and Information Security Handbook*; Elsevier: Amsterdam, The Netherlands, 2017; pp. 759–778.
7. Kunz, I.; Stephanow, P.; Banse, C. An Edge Framework for the Application of Privacy Enhancing Technologies in IoT Communications. In Proceedings of the ICC 2020—2020 IEEE International Conference on Communications (ICC), Dublin, Ireland, 7–11 June 2020; pp. 1–6. [CrossRef]
8. Alharbi, R.; Almagwashi, H. The Privacy Requirments for Wearable IoT Devices in Healthcare Domain. In Proceedings of the 2019 7th International Conference on Future Internet of Things and Cloud Workshops (FiCloudW), Istanbul, Turkey, 26–28 August 2019; pp. 18–25.
9. Aun, J.; Hurtado-Ram, D.; Porras-Díaz, L.; Irigoyen-Pen, B.; Rahmian, S.; Al-Khazraji, Y.; Soler-Garrido, J.; Kotsev, A. Evaluation and Utilisation of Privacy Enhancing Technologies—A Data Spaces Perspective. *Data Brief* **2024**, *55*, 110560. [CrossRef]
10. Chatterjee, A.; Aung, K.M.M. *Fully Homomorphic Encryption in Real World Applications*; Springer: Berlin/Heidelberg, Germany, 2019.
11. Li, J.J.; Lin, X.; Tang, C.; Lu, Y.Q.; Hu, X.; Zuo, E.; Li, H.; Ying, W.; Sun, Y.; Lai, L.L.; et al. Disruption of splicing-regulatory elements using CRISPR/Cas9 to rescue spinal muscular atrophy in human iPSCs and mice. *Natl. Sci. Rev.* **2020**, *7*, 92–101. [CrossRef] [PubMed]
12. Javed, I.T.; Alharbi, F.; Margaria, T.; Crespi, N.; Qureshi, K.N. PETchain: A Blockchain-Based Privacy Enhancing Technology. *IEEE Access* **2021**, *9*, 41129–41143. [CrossRef]
13. Garrido, G.M.; Schmidt, K.; Harth-Kitzerow, C.; Klepsch, J.; Luckow, A.; Matthes, F. Exploring privacy-enhancing technologies in the automotive value chain. In Proceedings of the 2021 IEEE International Conference on Big Data (Big Data), Orlando, FL, USA, 15–18 December 2021; pp. 1265–1272. [CrossRef]
14. Garrido, G.M.; Sedlmeir, J.; Uludağ, Ö.; Alaoui, I.S.; Luckow, A.; Matthes, F. Revealing the landscape of privacy-enhancing technologies in the context of data markets for the IoT: A systematic literature review. *J. Netw. Comput. Appl.* **2022**, *207*, 103465. [CrossRef]
15. Ustundag Soykan, E.; Karaçay, L.; Karakoç, F.; Tomur, E. A Survey and Guideline on Privacy Enhancing Technologies for Collaborative Machine Learning. *IEEE Access* **2022**, *10*, 97495–97519. [CrossRef]
16. Nguyen, D.C.; Pham, Q.V.; Pathirana, P.N.; Ding, M.; Seneviratne, A.; Lin, Z.; Dobre, O.; Hwang, W.J. Federated Learning for Smart Healthcare: A Survey. *ACM Comput. Surv.* **2022**, *55*, 1–37. [CrossRef]

17. Liu, Q.; Zhou, F.; Chen, H. Secure medical data on cloud storage via DNA homomorphic encryption technique. *Phys. Commun.* **2024**, *64*, 102295. [CrossRef]
18. Waheed, N.; Khan, F.; Mastorakis, S.; Jan, M.A.; Alalmaie, A.Z.; Nanda, P. Privacy-Enhanced Living: A Local Differential Privacy Approach to Secure Smart Home Data. In Proceedings of the 2023 IEEE International Conference on Omni-layer Intelligent Systems (COINS), Berlin, Germany, 23–25 July 2023; pp. 1–6.
19. Sahlabadi, M.; Shukur, Z.; Muniyandi, R.C.; SaberiKamarposhti, M. GDP: Group-Based Differential Privacy Framework for Secure Process Mining in the Internet of Medical Things. In Proceedings of the 2023 International Conference on Electrical Engineering and Informatics (ICEEI), Bandung, Indonesia, 10–11 October 2023; pp. 1–6.
20. Rane, J.; Mallick, S.; Kaya, O.; Rane, N. Federated learning for edge artificial intelligence: Enhancing security, robustness, privacy, personalization, and blockchain integration in IoT. In *Future Research Opportunities for Artificial Intelligence in Industry 4.0 and 5.0*; Deep Science Publishing: Erzurum, Turkey, 2024; Volume 5, pp. 2–94.
21. Zhang, C.; Xie, Y.; Bai, H.; Yu, B.; Li, W.; Gao, Y. A survey on federated learning. *Knowl.-Based Syst.* **2021**, *216*, 106775. [CrossRef]
22. Panigrahi, M.; Bharti, S.; Sharma, A. Federated Learning for Beginners: Types, Simulation Environments, and Open Challenges. In Proceedings of the 2023 International Conference on Computer, Electronics & Electrical Engineering & Their Applications (IC2E3), Srinagar Garhwal, India, 8–9 June 2023; pp. 1–6.
23. Oldenhof, M.; Ács, G.; Pejó, B.; Schuffenhauer, A.; Holway, N.; Sturm, N.; Dieckmann, A.; Fortmeier, O.; Boniface, E.; Mayer, C.; et al. Industry-scale orchestrated federated learning for drug discovery. In Proceedings of the AAAI Conference on Artificial Intelligence, Washington, DC, USA, 7–14 February 2023; Volume 37, pp. 15576–15584.
24. Pedrouzo-Ulloa, A.; Ramon, J.; Péerez-González, F.; Lilova, S.; Duflot, P.; Chihani, Z.; Gentili, N.; Ulivi, P.; Hoque, M.A.; Mukammel, T.; et al. Introducing the TRUMPET project: TRUstworthy Multi-site Privacy Enhancing Technologies. In Proceedings of the 2023 IEEE International Conference on Cyber Security and Resilience (CSR), Venice, Italy, 31 July–2 August 2023; pp. 604–611.
25. *ISO/IEC 29100:2024*; Information Technology—Security Techniques—Privacy Framework. ISO: Geneva, Switzerland, 2024.
26. Li, J.; Meng, Y.; Ma, L.; Du, S.; Zhu, H.; Pei, Q.; Shen, X. A Federated Learning Based Privacy-Preserving Smart Healthcare System. *IEEE Trans. Ind. Inform.* **2022**, *18*, 2021–2031. [CrossRef]
27. Terhörst, P.; Huber, M.; Damer, N.; Rot, P.; Kirchbuchner, F.; Struc, V.; Kuijper, A. Privacy Evaluation Protocols for the Evaluation of Soft-Biometric Privacy-Enhancing Technologies. In Proceedings of the 2020 International Conference of the Biometrics Special Interest Group (BIOSIG), Darmstadt, Germany, 16–18 September 2020; pp. 1–5.
28. Haddad, Z. Enhancing privacy and security in 5G networks with an anonymous handover protocol based on Blockchain and Zero Knowledge Proof. *Comput. Netw.* **2024**, *250*, 110544. [CrossRef]
29. Gatha; Chauhan, R.; Singh, D. Ensuring Privacy-Aware Data Release: An Analysis of Applicability of Privacy Enhancing Techniques to Real-World Datasets. In Proceedings of the 2020 8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), Noida, India, 4–5 June 2020; pp. 883–887.
30. Liu, R.; Zeighami, S.; Lin, H.; Shahabi, C.; Cao, Y.; Takagi, S.; Konishi, Y.; Yoshikawa, M.; Xiong, L. Supporting Pandemic Preparedness with Privacy Enhancing Technology. In Proceedings of the 2023 5th IEEE International Conference on Trust, Privacy and Security in Intelligent Systems and Applications (TPS-ISA), Atlanta, GA, USA, 1–4 November 2023; pp. 34–43.
31. Miller, J.; Chattopadhyay, A. Integrating Differential Privacy in Modern Database Curriculum. In Proceedings of the 2024 IEEE Integrated STEM Education Conference (ISEC), Princeton, NJ, USA, 9 March 2024; pp. 1–6.
32. Li, C.; Palanisamy, B. Privacy in Internet of Things: From Principles to Technologies. *IEEE Internet Things J.* **2019**, *6*, 488–505. [CrossRef]
33. Rai, H.M.; Shukla, K.K.; Tightiz, L.; Padmanaban, S. Enhancing data security and privacy in energy applications: Integrating IoT and blockchain technologies. *Heliyon* **2024**, *10*, e38917. [CrossRef] [PubMed]
34. Tandon, R.; Gupta, P. Chapter 10—Security and privacy challenges in healthcare using Internet of Things. In *IoT-Based Data Analytics for the Healthcare Industry*; Singh, S.K., Singh, R.S., Pandey, A.K., Udmale, S.S., Chaudhary, A., Eds.; Intelligent Data-Centric Systems; Academic Press: Cambridge, MA, USA, 2021; pp. 149–165.
35. Parihar, A.; Prajapati, J.B.; Prajapati, B.G.; Trambadiya, B.; Thakkar, A.; Engineer, P. Role of IOT in healthcare: Applications, security & privacy concerns. *Intell. Pharm.* **2024**, *2*, 707–714.
36. Louassef, B.R.; Chikouche, N. Privacy preservation in healthcare systems. In Proceedings of the 2021 International Conference on Artificial Intelligence for Cyber Security Systems and Privacy (AI-CSP), El Oued, Algeria, 20–21 November 2021; pp. 1–6.
37. Zhang, R.; Xue, R.; Liu, L. Security and Privacy for Healthcare Blockchains. *IEEE Trans. Serv. Comput.* **2022**, *15*, 3668–3686. [CrossRef]
38. Thapa, C.; Camtepe, S. Precision health data: Requirements, challenges and existing techniques for data security and privacy. *Comput. Biol. Med.* **2021**, *129*, 104130. [CrossRef]
39. Zhan, Y.; Meng, Y.; Zhou, L.; Zhu, H. Vetting Privacy Policies in VR: A Data Minimization Principle Perspective. In Proceedings of the IEEE INFOCOM 2023—IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS), Hoboken, NJ, USA, 20–20 May 2023; pp. 1–2.

40. Lax, G.; Canino, A.; Musarella, L. A Blockchain-Based Approach for Certifying Information at Different Levels of Granularity According to the Data Minimization Principle of the GDPR. In Proceedings of the 2024 8th International Conference on Computer, Software and Modeling (ICCSM), Paris, France, 4–6 July 2024; pp. 31–35.
41. Goldsteen, A.; Ezov, G.; Shmelkin, R.; Moffie, M.; Farkash, A. Data minimization for GDPR compliance in machine learning models. *AI Ethics* **2022**, *2*, 477–491. [CrossRef]
42. Eichinger, T.; Küpper, A. On data minimization and anonymity in pervasive mobile-to-mobile recommender systems. *Pervasive Mob. Comput.* **2024**, *103*, 101951. [CrossRef]
43. Mukta, R.; Paik, H.-y.; Lu, Q.; Kanhere, S.S. A survey of data minimisation techniques in blockchain-based healthcare. *Comput. Netw.* **2022**, *205*, 108766. [CrossRef]
44. Senarath, A.; Arachchilage, N.A.G. A data minimization model for embedding privacy into software systems. *Comput. Secur.* **2019**, *87*, 101605. [CrossRef]
45. Chen, Z.; Liao, G.; Ma, Q.; Chen, X. Adaptive Privacy Budget Allocation in Federated Learning: A Multi-Agent Reinforcement Learning Approach. In Proceedings of the ICC 2024—IEEE International Conference on Communications, Denver, CO, USA, 9–13 June 2024; pp. 5166–5171.
46. Nguyen, D.C.; Ding, M.; Pathirana, P.N.; Seneviratne, A.; Li, J.; Vincent Poor, H. Federated Learning for Internet of Things: A Comprehensive Survey. *IEEE Commun. Surv. Tutor.* **2021**, *23*, 1622–1658. [CrossRef]
47. Chen, S.; Huang, Y. A privacy-preserving federated learning approach for airline upgrade optimization. *J. Air Transp. Manag.* **2025**, *122*, 102693. [CrossRef]
48. Korkmaz, A.; Alhonainy, A.; Rao, P. An Evaluation of Federated Learning Techniques for Secure and Privacy-Preserving Machine Learning on Medical Datasets. In Proceedings of the 2022 IEEE Applied Imagery Pattern Recognition Workshop (AIPR), Washington, DC, USA, 11–13 October 2022; pp. 1–7. [CrossRef]
49. Liu, B.; Lv, N.; Guo, Y.; Li, Y. Recent advances on federated learning: A systematic survey. *Neurocomputing* **2024**, *597*, 128019. [CrossRef]
50. Abaoud, M.; Almuqrin, M.A.; Khan, M.F. Advancing Federated Learning Through Novel Mechanism for Privacy Preservation in Healthcare Applications. *IEEE Access* **2023**, *11*, 83562–83579. [CrossRef]
51. Zeng, R.; Mi, B.; Huang, D. A Federated Learning Framework Based on CSP Homomorphic Encryption. In Proceedings of the 2023 IEEE 12th Data Driven Control and Learning Systems Conference (DDCLS), Xiangtan, China, 12–14 May 2023; pp. 196–201. [CrossRef]
52. Tyagi, S.; Rajput, I.S.; Pandey, R. Federated learning: Applications, Security hazards and Defense measures. In Proceedings of the 2023 International Conference on Device Intelligence, Computing and Communication Technologies, (DICCT), Dehradun, India, 17–18 March 2023; pp. 477–482.
53. Ajao, A.; Jonathan, O.; Adetiba, E. The Applications of Federated Learning Algorithm in the Federated Cloud Environment: A Systematic Review. In Proceedings of the 2024 International Conference on Science, Engineering and Business for Driving Sustainable Development Goals (SEB4SDG), Omu-Aran, Nigeria, 2–4 April 2024; pp. 1–15.
54. Li, Z.; He, S.; Chaturvedi, P.; Hoang, T.H.; Ryu, M.; Huerta, E.A.; Kindratenko, V.; Fuhrman, J.; Giger, M.; Chard, R.; et al. APPFLx: Providing Privacy-Preserving Cross-Silo Federated Learning as a Service. In Proceedings of the 2023 IEEE 19th International Conference on e-Science (e-Science), Limassol, Cyprus, 9–13 October 2023; pp. 1–4.
55. Oskoui, S.E.; Retford, M.; Forde, E.; Barnes, R.; Hunter, K.J.; Wozencraft, A.; Thompson, S.; Orton, C.; Ford, D.; Heys, S.; et al. Developing a prototype for federated analysis to enhance privacy and enable trustworthy access to COVID-19 research data. *Int. J. Med. Inform.* **2024**, *195*, 105708. [CrossRef] [PubMed]
56. R, R.C.; Harshini, P.S.; N, T.; R, C.T.; Srinivas, D.B. A Multi-Stage Partial Homomorphic Encryption Scheme for Secure Data Processing in Cloud Computing. In Proceedings of the 2023 2nd International Conference on Edge Computing and Applications (ICECAA), Namakkal, India, 19–21 July 2023; pp. 58–62.
57. Curzon, J.; Almehmadi, A.; El-Khatib, K. A survey of privacy enhancing technologies for smart cities. *Pervasive Mob. Comput.* **2019**, *55*, 76–95. [CrossRef]
58. Xiong, J.; Chen, J.; Lin, J.; Jiao, D.; Liu, H. Enhancing privacy-preserving machine learning with self-learnable activation functions in fully homomorphic encryption. *J. Inf. Secur. Appl.* **2024**, *86*, 103887. [CrossRef]
59. Ni, C.; Cang, L.S.; Gope, P.; Min, G. Data anonymization evaluation for big data and IoT environment. *Inf. Sci.* **2022**, *605*, 381–392. [CrossRef]
60. Montenegro, H.; Cardoso, J.S. Anonymizing medical case-based explanations through disentanglement. *Med. Image Anal.* **2024**, *95*, 103209. [CrossRef] [PubMed]
61. Madan, S.; Goswami, D.P. An Extensive Study on Statistical Data Anonymization Algorithms. In Proceedings of the 2018 3rd International Conference and Workshops on Recent Advances and Innovations in Engineering (ICRAIE), Jaipur, India, 22–25 November 2018; pp. 1–5.

62. Majeed, A.; Lee, S. Anonymization Techniques for Privacy Preserving Data Publishing: A Comprehensive Survey. *IEEE Access* **2021**, *9*, 8512–8545. [CrossRef]
63. Ali, M.S.; Ahsan, M.M.; Tasnim, L.; Afrin, S.; Biswas, K.; Hossain, M.M.; Ahmed, M.M.; Hashan, R.; Islam, M.K.; Raman, S. Federated Learning in Healthcare: Model Misconducts, Security, Challenges, Applications, and Future Research Directions–A Systematic Review. *arXiv* **2024**, arXiv:2405.13832.