

Privacy-Enhancing Technologies in Collaborative Healthcare Analysis

Manar Alnasser^{*†}, and Shancang Li^{*‡} ^{*}School of Computer Science and Informatics, Cardiff University, Cardiff CF10 3AT, UK [†]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 the 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.

Index Terms—privacy, privacy enhancing technologies, healthcare, collaboration analysis

I. 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. 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. 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. Based on Statista, the healthcare industry is considered one of the most vulnerable to cybercrime. In 2023, it remained the most targeted by cyberattacks 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.

II. 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.

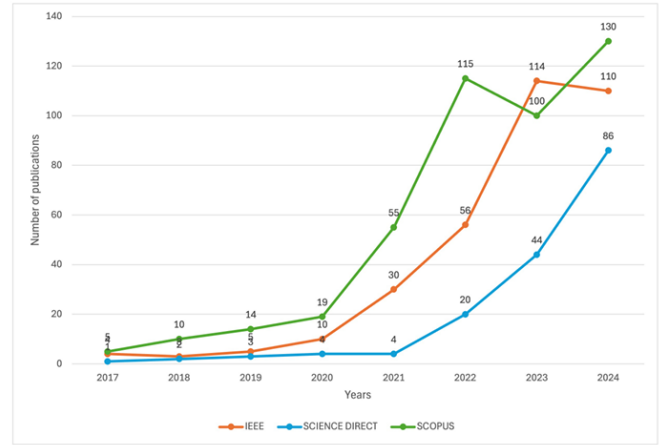


Fig. 1. The number of publications in PETs in the healthcare industry. Please insert the actual image of the chart here.

The work aims to provide a comprehensive overview of the use of PETs in the healthcare industry by reviewing 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?.

III. 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. This section will highlight privacy challenges and requirements in the healthcare industry.

TABLE I: 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.
[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.

Continued on next page

Table I continued from previous page

Ref.	Methods	Strengths	Limitations
[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.	This paper has highlighted how PETs can help overcome the challenge of fragmented data, enabling secure collaboration while adhering to strict privacy regulations like HIPAA and GDPR. Discover more effective features to represent the characteristics of ADs. The survey examined key PETs, including Data Minimization, Federated Learning, and Anonymization, and discussed the specific privacy requirements and challenges of the healthcare industry. These insights provide a foundational understanding for researchers and practitioners, paving the way for the development of more effective privacy-preserving frameworks in the future. The findings suggest that the adoption of these technologies is essential for both improving patient care and maintaining the integrity of sensitive health data.
[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. The lack of interoperability.

IV. 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**.

A. 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**. These regulations dictate that only required data be collected to fulfil a certain purpose. In the literature, data minimisation is commonly defined with phrases 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.

B. Federated Learning (FL)

A federated learning has emerged as a powerful machine learning approach aims to protecting the privacy of data. It is based on a principle of training machine learning models on decentralized entities holding local data without sharing them. 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.

V. CONCLUSION

This report has provided a comprehensive overview of privacy-enhancing technologies (PETs) within the context of collaborative healthcare analysis. As the digitization of patient data continues, the need for robust privacy solutions is more

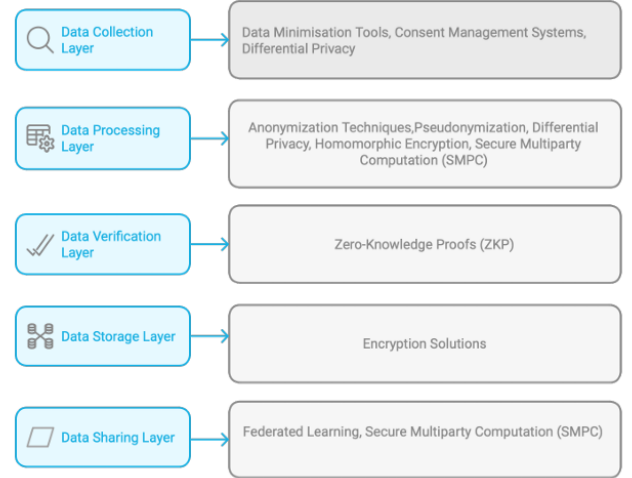


Fig. 2. The taxonomy of key PETs based on their usage. (a) Essential Privacy requirements; (b) Key challenges. Please insert the actual image of the charts here.