International Journal of Engineering Research and Science & Technology

PRIVACY PRESERVING HEALTH CARE DATA SHARING WITH DATA MINING

DARAM KUSUMA KUMARI [1], CHALAMALASETTY MANISH KUMAR [2], YAMSANI PRAVEEN KUMAR [3], EDIGI SAI CHARAN [4], JANGILI GOUTHAM [5]

[1]Assistant Professor, Dept. of CSE, Guru Nanak Institutions Technical Campus**,** Ibrahimpatnam, Telangana, India

[2, 3, 4, 5] B.Tech, Dept. of CSE, Guru Nanak Institutions Technical Campus**,** Ibrahimpatnam, Telangana, India

ABSTRACT— The growing importance of data in healthcare has heightened the need for privacy-preserving data sharing, particularly in systems where health records are distributed across a databases. Despite widespread recognition of the benefits of data sharing for both research and patient care, concerns about privacy and security remain a major barrier. This study explores current attitudes toward data sharing healthcare professionals across clinical and non-clinical roles. By combining descriptive statistics and data mining techniques, we assessed trust in existing privacy-preserving tools—such as data anonymization, encryption, and access control mechanisms—and evaluated openness to adopting proposed advanced solutions, including differential privacy, and secure multi-party. Findings reveal a cautious with many professionals showing readiness, primarily motivated by the potential to improve patient outcomes rather than purely research interests. These insights contribute to the development of targeted policies and innovative frameworks aimed at enabling secure, privacy-conscious data sharing in the healthcare sector.

Index Terms – Privacy Preservation, Healthcare Data Sharing, Data Mining, Electronic Health Records, Data Security, Healthcare Analytics, Cryptography, Access Control, Secure Data Sharing, Cybersecurity.

I. INTRODUCTION

The rise of big data – large and complex data sets growing at an exponential rate – has brought increased urgency to the privacy-preserving data sharing among organisations for various reasons including the mining of the combined data for knowledge discovery. This is an important topic for healthcare,

where vast amounts of patient records. are generated and maintained privately by an ever-growing number of organisations. Significantly, this data has the potential to unlock knowledge and hidden patterns through data mining, which can help to solve real-world problems. Healthcare companies and hospitals securely store confidential medical records encrypted in data centres. With the rise of big data and a more informed consumer, public views on privacy are changing, leading to conflicting attitudes between privacy protection and public benefit In Australia, advances in cloud computing technology in

healthcare such as My Health Record the electronic medical record platform managed by the Australian government, have opened the door to many potential benefits including the sharing of clinical data to improve healthcare.

While most consumers and healthcare professionals see the benefit of data-sharing for research purposes concerns about privacy and security continue to persist. In the Australian health industry, these concerns have influenced the adoption of a complex approvals process intended to protect personal information, but the process is often inefficient at allowing easy data-sharing in practice. Despite current concerns about data privacy, there is a lack of research on whether technological solutions for privacy preserving data sharing would be embraced by the Australian healthcare sector should these concerns be adequately addressed. This study aims to contribute by gaining an understanding of current attitudes and levels of trust in technological solutions for privacy-preserving data sharing in healthcare and to understand if technological solutions would possibly be adopted.

The primary goal of this research is to examine current attitudes toward data sharing in the Australian healthcare sector and to evaluate current levels of trust and concern among healthcare professionals and organisations. The study also explores attitudes towards the adoption of technological solutions to privacy-preserving data sharing.

II. RELATED WORK

A. A privacy- preserving and redactable healthcare blockchain system

Blockchain as an open and immutable ledger is being posited as the next frontier in healthcare that will help solve the industry's interoperability challenges. However, immutability in processing personal data is

no longer legal since the General Data Protection Regulation (GDPR) requires the "right to be forgotten" as a critical data subject right. To observe such data regulation, it is desirable to build a healthcare blockchain with data redaction in a controlled way. Moreover, electronic health records (EHRs) usually are sensitive and the conventional blockchain lacks systematic and formal security analysis of data confidentiality, especially in the multi-user setting. Furthermore, EHRs are typically helpful in medical research for predicting epidemic diseases and valuable in insurance agencies making business plans. Hence, in healthcare blockchain systems, data confidentiality and flexible key distribution have become the most challenging issues that should be urgently resolved. In this article, we propose a privacy-preserving and redactable healthcare blockchain system (PRHBS). Our solution offers fine-grained block-level data reduction and secure data sharing with flexible key distribution mechanisms. We give the formal definition and security models of PRHBS, and propose a generic construction based on trapdoor-based chameleon-hash function, attribute-based encryption, and puncturable encryption. We present formal security analysis and give an instantiation based on our proposed generic construction. The comprehensive comparison and experimental simulation demonstrate that our implementation exhibits comparable performance, while surpassing the most relevant solutions in terms of functionality.

B. Privacy-preserving and secure industrial big data analytics: A survey and the research framework

The development of the Industrial Internet will generate a large amount of valuable data, known as industrial big data (IBD). By mining and utilizing IBD, enterprises can improve production efficiency, reduce costs and risks, optimize management processes, and innovate services and business models. However, IBD comes from various institutions in all walks of life and has features such as multisource, heterogeneity, and multimodality. And data sharing and trading (DS&T) occur in the Industrial Internet environment without mutual trust. These characteristics pose new challenges to analytics methods and privacy and security protection technologies. Therefore, this article aims to provide references for privacy-preserving and secure industrial big data analytics (IBDA) from three perspectives: 1) research framework; 2) platform architecture; and 3) key technologies. First,

we review the current state of research on theories and technologies related to IBDA. Then, we reveal three challenges to secure and efficient IBDA. We take the analytics and utilization of IBD as systematic engineering, propose the research framework for privacy-preserving and secure IBDA, and point out the specific content to be studied. Further, we design the architecture of the IBDA platform with the idea of layering, including a function model, security architecture, and system architecture. Finally, detailed research proposals and potential technologies for IBD analytics and utilization are presented from three aspects: 1) data fusion and analytics; 2) data privacy and security protection; and 3) blockchain.

C. Gathering data for decisions: Best practice use of primary care electronic records for research

In Australia, there is limited use of primary health care data for research and for data linkage between health care settings. This puts Australia behind many developed countries. In addition, without use of primary health care data for research, knowledge about patients' journeys through the health care system is limited. There is growing momentum to establish "big data" repositories of primary care clinical data to enable data linkage, primary care and population health research, and quality assurance activities. However, little research has been conducted on the general public's and practitioners' concerns about secondary use of electronic health records in Australia. International studies have identified barriers to use of general practice patient records for research. These include legal, technical, ethical, social and resource-related issues. Examples include concerns about privacy protection, data security, data custodians and the motives for collecting data, as well as a lack of incentives for general practitioners to share data. Addressing barriers may help define good practices for appropriate use of health data for research. Any model for general practice data sharing for research should be underpinned by transparency and a strong legal, ethical, governance and data security framework. Mechanisms to collect electronic medical records in ethical, secure and privacy-controlled ways are available. Before the potential benefits of health-related data research can be realised, Australians should be well informed of the risks and benefits so that the necessary social licence can be generated to support such endeavours.

III. PROPOSED METHODOLOGY

This advanced form of encryption allows computations to be performed directly on encrypted data without decrypting it first.

Though computationally intensive, it provides a way to analyse data securely without ever exposing the raw data. This ensures that raw data never leaves its source, significantly reducing the risk of data breaches. Technological solutions like encryption provide essential tools that can ensure patient privacy and security but currently restricts data mining activities on the data. Typical encryption techniques provide very limited opportunities to perform data mining on encrypted data prior to decryption

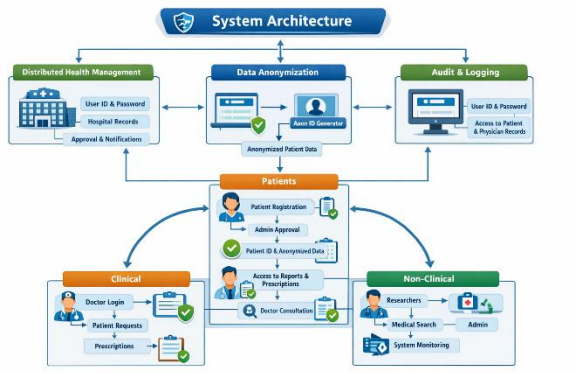
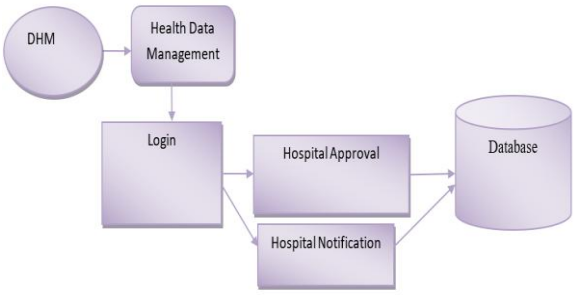


Fig. 1: System Overview

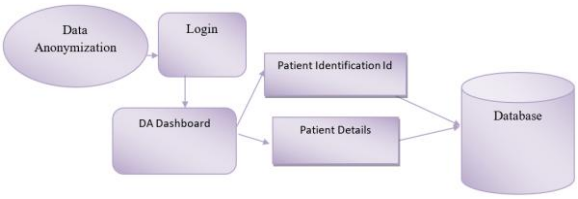
1. Distributed Health Management

The Distributed Health module ensures secure and controlled access to hospital data by requiring a valid user ID and password. Each hospital is registered in the system with essential details, including identification information and login credentials. This authentication mechanism prevents unauthorized access and protects sensitive health records. Once registered, hospitals can access distributed health data stored across multiple locations, enabling efficient data sharing while maintaining privacy. Before accessing or contributing data, hospitals must receive approval from the distributed health authority. This approval process verifies the authenticity of the hospital and ensures compliance with healthcare regulations. After approval, hospitals can publish and access disseminated health notices, such as updates on patient care, medical alerts, or administrative information. The distributed health notice mechanism allows timely and secure communication among authorized hospitals. Overall, this module improves data availability, enhances collaboration between healthcare institutions, and ensures confidentiality, integrity, and reliability of distributed health information across the network.



2. Data Anonymization

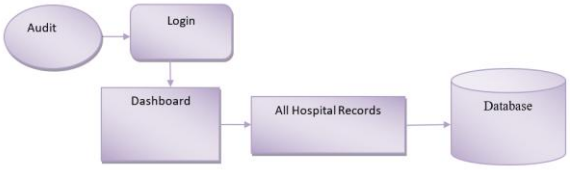
The Data Anonymization module ensures the privacy and security of patient information by controlling access through a secure user ID and password login. Only authorized users can access the anonymization system, preventing unauthorized handling of sensitive medical data. During the data anonymization process, a unique patient identity number is generated to replace personally identifiable information such as name, address, and contact details. This identity number allows patient records to be processed, analysed, and shared without revealing the patient's real identity. All patient-related information is included in the anonymization module, but sensitive attributes are masked or transformed to maintain confidentiality. By separating identity from medical data, the system reduces the risk of data breaches and supports compliance with healthcare data protection regulations. The Data Anonymization module enables secure data usage for research, analysis, and distributed healthcare systems while preserving patient privacy, data integrity, and trust within the healthcare ecosystem.



3. Audit & Logging Module

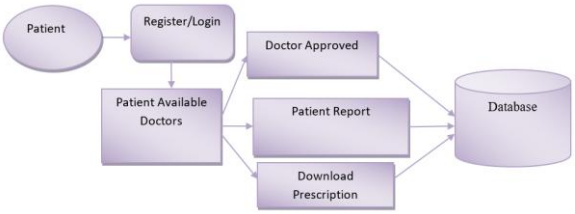
The Audit and Logging module plays a critical role in maintaining transparency, accountability, and security within the healthcare system. Access to the audit module is strictly controlled through a secure user ID and password login, ensuring that only authorized audit personnel can view sensitive records. Once authenticated, the audit system has access to all hospital records, including patient data, physician information, and system activity logs. Every action performed within the healthcare platform—such as data access, updates, approvals, and downloads—is recorded with timestamps and user details. This comprehensive logging mechanism helps detect unauthorized activities, data misuse, or policy violations. Audit logs also support regulatory compliance, legal investigations, and internal reviews by providing a complete

history of system operations. By monitoring interactions across hospitals, patients, and physicians, the Audit and Logging module enhances data integrity, strengthens security controls, and ensures trust and accountability throughout the distributed healthcare environment.



4. Patients

In the healthcare system, the patient begins by registering with personal and medical information. After registration, the patient must obtain authorization from the administrator to ensure the legitimacy and security of the account. Once the administrator grants approval, a unique patient ID is generated, and the patient's data is anonymized before login to protect privacy. This anonymization process ensures that sensitive personal details are securely masked. A physician is assigned to the patient, enabling proper medical consultation and treatment. The physician reviews the patient's condition and provides approval for medical reports and prescriptions. The patient can upload and access medical reports through the system and view authorized healthcare data. Additionally, the patient has access to prescription bills related to their treatment. After approval, the patient can securely download the prescription for future reference or medication purchases. This process ensures privacy, security, and efficient healthcare management.



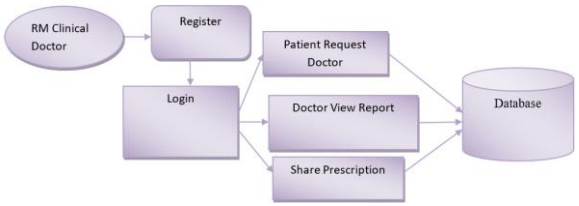
5.Role Management

The role management has a categories into a two types they are

1.Clinicical

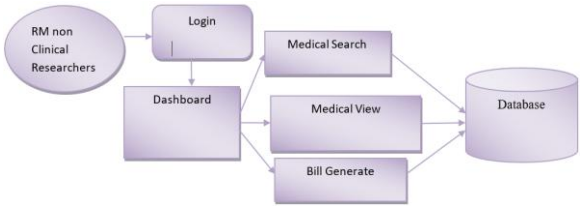
The Clinical module manages doctor-related information and ensures secure interaction between physicians and patients. Each doctor is registered in the system with complete professional details stored in the clinical database. Before accessing the system, the doctor must obtain permission from the hospital, which verifies the physician's credentials and authorizes system usage. Once approved, the physician logs in securely using a user ID and password. Patients can submit medical requests to the physician through the

platform, enabling efficient communication and consultation. Upon receiving a request, the physician reviews the patient's medical report and examines the provided health data carefully. Based on the analysis, the doctor prepares an appropriate diagnosis and issues a prescription for the patient. The prescription is securely shared with the patient through the system, ensuring accuracy and confidentiality. This clinical workflow improves coordination between hospitals, doctors, and patients while maintaining data security, accountability, and effective healthcare delivery.



2.Non Clinicical

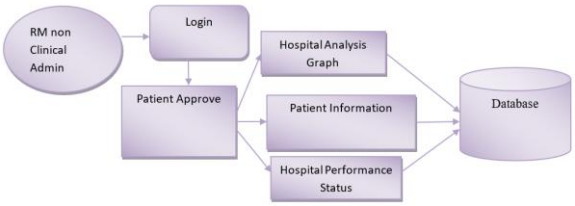
Researchers play an important role in analysing medical data and supporting healthcare decision-making. Each researcher is provided with a secure user ID and password to access the medical system. After logging in, researchers can conduct medical searches related to medications, treatments, and patient outcomes. Their access allows them to study medication effectiveness, dosage patterns, and associated charges for patient medications. From a medical perspective, researchers analyze anonymized patient data to generate insights that can improve treatment quality, reduce costs, and support medical innovation. Their activities are restricted to research purposes, ensuring patient privacy and data security.



3. Admin

The Admin module controls and monitors overall system operations. The administrator logs in securely using a user ID and password. One of the admin's key responsibilities is requesting and approving patient data access to maintain compliance and security. The administrator can view graphical reports generated from hospital data, including hospital performance statistics and patient performance metrics selected by specific hospitals. These visual analytics help assess service quality, efficiency, and outcomes. The admin also monitors overall hospital performance status, enabling informed decisions, system optimization, and effective healthcare management.

International Journal of Engineering Research and Science & Technology



IV. EXPERIMENTAL RESULTS AND ANALYSIS

Experimental results demonstrated that the system successfully enabled secure sharing and analysis of healthcare data while protecting sensitive patient information from unauthorized access. The privacy-preserving mechanisms maintained data confidentiality without significantly affecting the performance of data mining operations. The analysis showed that the integration of data mining techniques improved the extraction of meaningful healthcare insights and supported efficient decision-making in medical applications. Furthermore, the framework achieved low communication and processing overhead while maintaining scalability with increasing numbers of users and medical records.

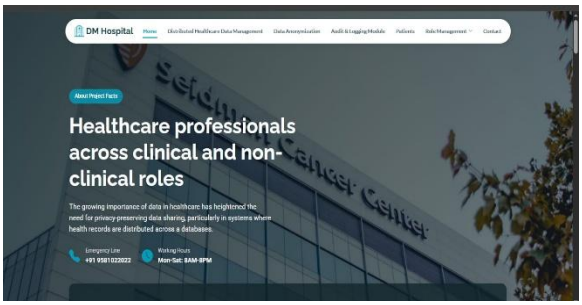


Fig. 2: Homepage

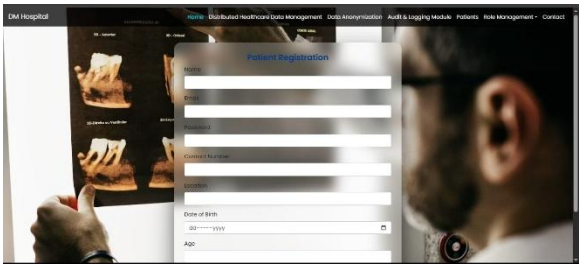


Fig. 3: Patient Registration

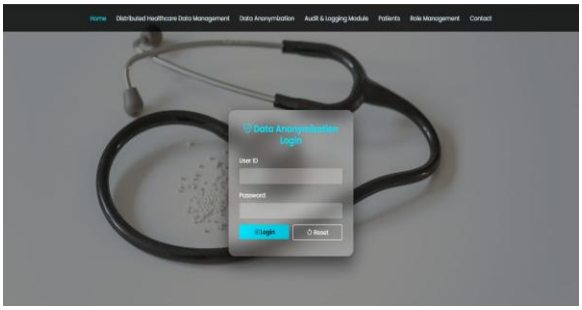


Fig. 4: Login

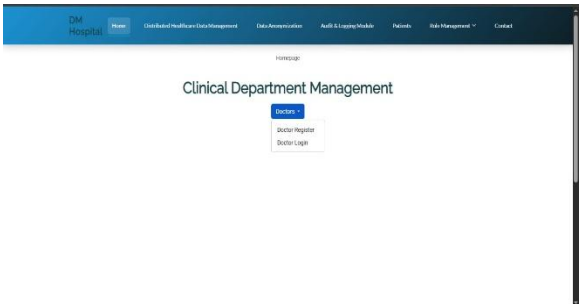


Fig. 5: Clinical Department

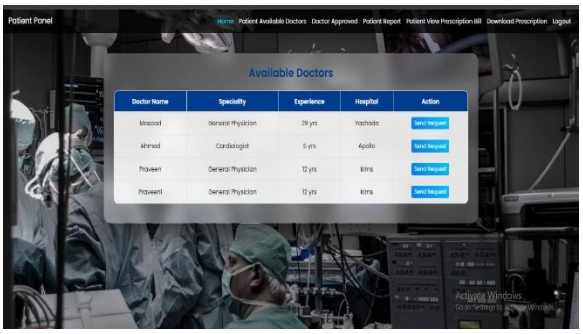


Fig. 5: Available Doctors

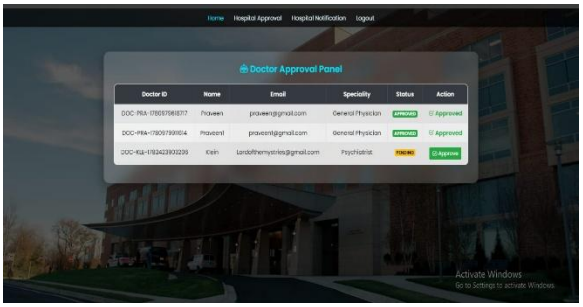


Fig. 6: Doctor Approval Panel

V. CONCLUSION

This study examined current attitudes toward data sharing in the Australian healthcare sector and identified level of trust and concern among healthcare professionals and organisations. We also explored the barriers and potential adoption of technological solutions to privacy-preserving data sharing, which has received limited attention in previous research. We have highlighted several limitations of this study, including the limiting effect of the sample size on measuring levels of trust and concern within some of the variables and the timing of the study possibly influencing some of the results related to barriers to adoption. The majority of respondents fell into the mid-aged bracket, and this also had the effect of some of the findings being centred on this age group. Future research should focus on expanding on the sample size, and having more data collected in different economic cycles would help provide more information about the barriers that exist. Additionally, we feel that a view of any contrasts in Australia's unique private-public healthcare practice split would be pertinent in future studies, especially with the context of cost being a major barrier.

REFERENCES

[1] S. Xu, J. Ning, X. Li, J. Yuan, X. Huang, and R. H. Deng, ''A privacypreserving and redactable healthcare blockchain system,'' IEEE Trans. Services Compute., vol. 17, no. 2, pp. 364–377, Mar. 2024.

[2] I. Lunden. (Oct. 2021). Triple Blind Secures $24M for a New Approach to Enterprise-level, Privacy-Preserving Data Sharing. TechCrunch. [Online]. Available: https://techcrunch.com/2021/10/18/triple blind-secures-24m for-a-new-approach-to-enterprise-level-privacy-preserving-data-sharing/



[3] L. Liu, J. Li, J. Lv, J. Wang, S. Zhao, and Q. Lu, ‘‘Privacy-preserving and secure industrial big data analytics: A survey and the research framework,’’ IEEE Internet Things J., vol. 11, no. 11, pp. 18976–18999, Jun. 2024.

[4] V. Xafis, ‘‘The acceptability of conducting data linkage research without obtaining consent: Lay people’s views and justifications,’’ BMC Med. Ethics, vol. 16, no. 1, pp. 1–16, Dec. 2015.

[5] Australian Digital Health Agency. (2021). My Health Record, 2021. Accessed: 2023. [Online]. Available: https://www.digitalhealth.gov.au/initiatives-and-programs/my-health-record

[6] R. Canaway, D. I. Boyle, J.-A.-E. Manski-Nankervis, J. Bell, J. S. Hocking, K. Clarke, M. Clark, J. M. Gunn, and J. D. Emery, ‘‘Gathering data for decisions: Best practice use of primary care electronic records for research,’’ Med. J. Aust., vol. 210, no. S6, pp. S12–S16, Apr. 2019.

AUTHORS Profile

Mrs. DARAM KUSUMA KUMARI received her B.Tech degree from JNTUH, Hyderabad, in 2009 and her M.Tech degree in Computer Science from JNTUH, Hyderabad, in 2015. She is currently pursuing a Ph.D. in Computer Science at Bharathiya Engineering Science and Technology IU University. She has been dedicated to the teaching profession for the past 9.5 years. Her research interests include Cybersecurity with Artificial Intelligence (AI). She is currently working as an Assistant Professor in the Department of Computer Science and Engineering (CSE) at Guru Nanak Institutions Technical Campus, Ibrahimpatnam, Telangana, India.



Mr. CHALAMALASETTY MANISH KUMAR pursuing B.Tech degree in Computer Science and Engineering at Guru Nanak Institutions Technical Campus affiliated to JNTUH in 2022-2026.



Mr. YAMSANI PRAVEEN KUMAR pursuing B.Tech degree in Computer Science and Engineering at Guru Nanak Institutions

Technical Campus affiliated to JNTUH in 2022-2026.



Mr. EDIGI SAI CHARAN pursuing B.Tech degree in Computer Science and Engineering at Guru Nanak Institutions Technical Campus affiliated to JNTUH in 2022-2026.



Mr. JANGILI GOUTHAM pursuing B.Tech degree in Computer Science and Engineering at Guru Nanak Institutions Technical Campus affiliated to JNTUH in 2022-2026.