IJNRD.ORG ISSN : 2456-4184
INTERNATIONAL JOURNAL OF NOVEL RESEARCH AND DEVELOPMENT (IJNRD) | IJNRD.ORG
An International Open Access, Peer-reviewed, Refereed Journal

Safeguarding Patient Privacy in the Big Data Era: Strategies and Technologies for Ensuring Confidentiality in Healthcare Systems

Chaithra S, Sanskar Dutta, Sarita Bhagwanrao Biradar, Sathvika D S, Yogesh Kumar

Assistant Professor, Department of Computer Science and Engineering, Sri Venkateshwara College of Engineering, Karnataka, India
B.E, Department of Computer Science and Engineering, Sri Venkateshwara College of Engineering, Karnataka, India
B.E, Department of Computer Science and Engineering, Sri Venkateshwara College of Engineering, Karnataka, India
B.E, Department of Computer Science and Engineering, Sri Venkateshwara College of Engineering, Karnataka, India
B.E, Department of Computer Science and Engineering, Sri Venkateshwara College of Engineering, Karnataka, India

Abstract

Big data has completely changed the healthcare sector by providing advantages including prompt patient treatment, early disease identification, and real-time monitoring. Still, challenges arise from data poachers compromising patient data privacy. This study delves into technologies for preserving patient privacy in healthcare, focusing on Electronic Health Records (EHRs). Anonymization techniques, cryptography, and data management frameworks are explored, with data management frameworks emerging as a more reliable method for security and privacy preservation. Data has changed due to big data. management across industries, with healthcare being a promising field for its application. Big healthcare Data could lead to improvements. patient outcomes and healthcare delivery, finding a balance between data utility, security, and privacy, however, can be difficult. Resolving privacy and security issues is essential to successfully utilizing big data in healthcare. In order to improve data security and privacy in the healthcare sector, this article examines security and privacy concerns in big data, assesses current solutions—particularly anonymization and encryption techniques—and proposes future research approaches.

Keywords: Big Data, Strategies and technologies, Security, Privacy, Anonymization, Encryption, confidentiality.

INTRODUCTION

The emergence of big data technology has resulted ina significant upheaval in the healthcare sector in recent years. The emergence of electronic health records (EHRs), wearable technology, genomic sequencing, and additional healthcare data sources has led to an unparalleled amount and diversity of data being produced and gathered. There is great promise for bettering patient outcomes, allocating resources more efficiently, and advancing medical research with this abundance of data, However, there are also grave concerns over the privacy and security of private patient information. Protecting patient privacy has developed into atop priority for regulatory agencies and healthcare institutions in the age of big data.The digitization and centralization of healthcare data have made it increasingly susceptible to unauthorized access, breaches, and misuse. Moreover, theinterconnected nature of healthcare systems and the proliferation of data-sharing initiatives further exacerbate the risks associated with patient data privacy.

This study intends to investigate the technologies and approaches that can be utilized to ensure patient information confidentiality in the setting of big data in healthcare systems. By examining the intricate problems brought about by the combination of big data and patient privacy, this research aims to identify workable solutions for risk mitigation and safe keeping of private medical records. The importance of this research resides in its ability to educate legislators, medical professionals, and tech developers on the best practices and important factors to protect patient privacy in the big data era.By elucidating the complexities of this evolving landscape and offering actionable insights, this study seeks to contribute to the ongoing discourse on privacy and security in healthcare, ultimately fostering a safer and more secure environment for the storage, transmission, and utilization of healthcare data.

1. OVERVIEW USE OF BIG DATA IN HEALTHCARE

The advent of analytics for big datahas caused a paradigm change in the healthcare sector in recent years. Massive amounts of organized and unstructured data produced by a variety of sources, such as wearable technology, medical imaging, genetic sequencing, and electronic health records (EHRs), are referred to as big data. Healthcare organizations face a multitude of issues and opportunities as a result of this abundance of data, which also makes it possible for them to make better decisions, extract insightful information, and improve patient outcomes. Big data's three distinguishing features—volume, variety, and velocity—also known as the three Vs—underline its revolutionary potential in the medical

field. Healthcare practitioners could glean insightful information from vastdatasets and deliver more individualized and effective care by utilizing advanced analytics approaches like machine learning and predictive modelling.

2. APPLICATIONS OF BIG DATA IN HEALTHCARE

Big data analytics has a large number of uses in healthcare, revolutionizing various aspects of the industry. Predictive analytics, for instance, enables healthcare organizations to forecast disease outbreaks, identify high-risk patient populations, and optimize treatment plans. Population health management leverages big data to monitor the health status of communities, identify trends, and allocate resources effectively. In drug discovery and development, big data analytics accelerates the identification of potential drug candidates, streamlines clinical trials, and improves drug safety. Additionally, precision medicine harnesses genomic and molecular data to tailor treatment approaches to individual patients, maximizing efficacy and minimizing adverse effects.

3.CHALLENGES AND LIMITATIONS

Despite potential benefits, there are challenges associated with using big data in healthcare. Sensitive information found in healthcare data must be shielded from unauthorized access and security breaches. therefore data privacy and security issues are very real. A major obstacle to the smooth integration and interchange of data between various healthcare environments and systems is interoperability. Ethical issues such as data ownership and patient consent hamper the proper use of healthcare data. Furthermore, because healthcare data is both reliable and of high quality, there may be differences in the validity and accuracy of analytics insights. Regulatory compliance adds another layer, particularly when it comes to laws like the Health Insurance Portability and Accountability Act (HIPAA) in the US and the General Data Protection Regulation (GDPR) in the EU.

4. TOOLS AND TECHNOLOGIES

Using big data analytics in healthcare is facilitated by a multitude of technologies and instruments. Large datasets can be processed and analysed across distributed computing clusters thanks to frameworks for distributed computing like Hadoop and Spark. Tools for reporting and data visualization offer user-friendly interfaces for analyzing and comprehending complicated medical data. Predictive modelling and pattern recognition activities are powered by machine learning algorithms, such as neural networks and decision trees. Methods for processing natural language (NLP) make it easier to draw conclusions from medical literature and unstructured clinical notes. Platforms for cloud computing provide scalable and affordable infrastructure for handling and storing medical data safely.

5. ETHICAL AND REGULATORY CONSIDERATIONS

Big data in healthcare has significant ethical ramifications that raise concerns about patient privacy, autonomy, and beneficence. Patient data protection must be given top priority in healthcare institutions, and data use must adhere to accountability, transparency, and fairness standards. Regulations like GDPR and HIPAA impose legal requirements on technology suppliers and healthcare providers by establishing standards for the gathering, storing, and use of healthcare data. Concerns like informed permission, data anonymization, and appropriate sharing of data with outside partiesare all covered by ethical considerations.

6. FUTURE DIRECTIONS AND OPPORTUNITIES

Big data analytics will eventuallylikely spur more advancements in patient care and healthcare delivery. Technological developments in artificial intelligence, namely in the domains of natural language processing and deep learning, have the potential to improve diagnosis precision, forecast treatment results, and customize interventions. A new era of precision medicine, where therapies are customized to individual genetic profiles, is anticipated to begin with the integration of genomic data into clinical practice. Proactive interventions and preventive care techniques are made possible by real-time analytics and continuous monitoring technologies. Using digital technologies and data-driven insights, patient engagement and participatory healthcare models enable people to actively manage their own health.

STRATEGIES AND TECHNIQUES FOR PROTECTING PRIVACY IN LARGE DATA

Here is a brief description of a few conventional techniques for protecting privacy in large data sets. While these procedures have historically been employed to protect patient privacy [43–45], their drawbacks have prompted the development of other approaches.

1. ENCRYPTION AND TOKENIZATION

Encryption and Tokenization are both cryptographic techniques used to protect sensitive data, including health information containing patient identifiers, from unauthorized access and breaches. Even though their functions are identical, theyoperate in different ways:

a. Encryption: The method of encoding data so that only people with permission can view itand decode it is known as encryption. It entails utilizing a secret key and an encryption method to transform plaintext data into ciphertext. Anyone without the necessary key to decrypt the ciphertext sees it as gibberish. Data at rest, or data that is kept, and data in transit, or data that is being communicated over networks, are both frequently protected with encryption. Two primary categories of encryption exist:

b. Symmetric Encryption: The same Keys are used for decrypting and encrypting data.in symmetric encryption. The sender and the recipient must first securely share this key. Both the Advanced Encryption Standard (AES) and the Data Encryption Standard (DES)are two examples of symmetric encryption methods.

c. Asymmetric Encryption*:* A symmetric encryption, also referred to as public-key cryptography, employs two keys: a private key for decryption and a public key for encryption. While the private key is kept confidential, the public key is freely shared. Only the matching private key can be used to decrypt messages that have been encrypted using the public key. Elliptic Curve Cryptography (ECC) and RSA (Rivest–Shamir–Adleman) are two examples of asymmetric encryption methods.

TOKENIZATION

Tokenization is the process of substituting sensitive data with non-sensitive placeholder values, called tokens, while preserving the data's original format and length. Unlike encryption, which transforms data into ciphertext, tokenization does not use mathematical algorithms to alter the data. Instead, it simply replaces the data with a token, which serves as a reference to the original data stored in a

secure tokenization system. The tokenization process typically involves the following steps:

A tokenization system generates a unique token for each piece of sensitive data, such as a credit card number or a patient identifier.

The sensitive data is securely stored in a centralized token vault or database.

When a transaction or query involving the sensitive data occurs, the token is used in place of the actual data.

Authorized users can retrieve the original data from the token vault using the corresponding token.

Payment processing, healthcare, and other sectors requiring the protection of sensitive data frequently use tokenization. Reducing the scope of compliance requirements (such PCI DSS for credit card data) and lowering the danger of data breaches are only a few benefits it provides. Other benefits include making data storage and retrieval simpler.

2. LIMITATIONS ONACCESS CONTROLS AND ROLE BASED PERMISSIONS

Gain access Within an organization, access to resources, systems, and data is regulated and managed through the use of role-based permissions and can Access controls. They guarantee that access is restricted to authorized usersparticular information or taking particular activities, and that only authorized people or institutions are able to do so. Let's examine each idea in more details:

a. Access Controls:

Healthcare companies manage enormous amounts of sensitive patient data in the age of big data, including genetic data, medical imaging results, and electronic health records (EHRs). In order to ensure that only people with permission, such administrative personnel and healthcare practitioners, may access this data, access controls are crucial. Organizations can reduce the possibility of data leaksand prevent unauthorized access to patient data by adopting access controls at many levels, such as logical, administrative, and physical restrictions.

b. Physical Access Controls: Physical access controls, such as biometric scanners and access badges, limit physical access to locations where data is stored, server rooms, and other sensitive areas where patient data is stored or processed.

c. Logical Access Controls: Logical access controls, such as passwords, multi-factor authentication, and access control lists (ACLs), regulate access to computer systems, databases, and systems for electronic health records (EHRs). Before using, users must authenticate themselves. patient data, and their access privileges are based on predefined rules and permissions.

d. Administrative Access Controls: Administrative access controls involve user account management, privilege management, and user provisioning processes. Healthcare organizations must carefully manage user accounts, roles, and permissions to ensure that only authorized personnel have access to patient data, and that access rights are granted based on job roles and responsibilities.

ROLE-BASED PERMISSIONS

In the age of big data, protecting patient privacy is especially important when it comes to role-based permissions, or role-based access control (RBAC). Users are granted permissions via RBAC according to their jobs or responsibilities within the healthcare company. This strategy reduces the possibility of illegal access or disclosure by guaranteeing that users Roles: In healthcare settings, roles may include physicians, nurses, administrative staff, IT personnel, and other healthcare professionals. Each role is associated with specific access rights and permissions based on the responsibilities and functions of the role. only have access to the patient data necessary for their specific roles.

a. Permissions: Permissions define the actions or operations that users with a particular role can perform on patient data. These actions may include viewing patient records, updating treatment plans, conducting medical tests, or generating reports.

b. Role Assignment: Role assignment involves assigning users to roles based on their job roles, responsibilities, and qualifications. Users inherit the permissions associated with their assigned roles, ensuring that they have access only to the patient data necessary for their job functions.

c. Role Activation: Role activation determines when users assume specific roles and gain access to the corresponding permissions. Role activation may occur upon user authentication, user request, or other triggering events.

By implementing access controls and role-based permissions effectively, healthcare organizations can safeguard patient privacy in the age of big data. These steps aid in ensuring thatpatient data is accessed only by authorized individuals for legitimate purposes, while unauthorized access is prevented or detected promptly. Additionally, access controls and role-based permissions support compliance with privacy regulations such as HIPAA (Health Insurance Portability and Accountability Act) and GDPR (General Data Protection Regulation), which mandate the protection of patient privacy and security in healthcare environments.

3. AUDIT TRAILS AND MONITORING

Audit trails and monitoring are essential components of cybersecurity and data governance practices. They serve to track and record activities, events, and changes within systems, networks, and applications. Here's an explanation of each concept:

AUDIT TRAILS

Audit trails are sequential logs of events and actions that take place on a network or system. Audit trails offer a thorough history of interactions with patient data in healthcare settings, including who accessed the data, what was done, when it was done, and where it came from. When it comes to tracking and monitoring patient information access, identifying unauthorized or suspicious activity, and assisting with investigations in the event of security incidents or breaches, audit trails are an invaluable resource.

a. Recording Access Events: Audit trails record access events related to patient data, such as logins, queries, modifications, and deletions. Each access event is logged along with relevant details, including the user's identity, timestamp, and the specific data accessed.

b. Maintaining Data Integrity: Audit trails help maintain the integrity of patient data by providing a verifiable record of all actions taken on the data. Any unauthorized or suspicious activities can be identified and investigated promptly, helping to guarantee the preservation of patient data accurate, reliable, and trustworthy.

c. Supporting Compliance Requirements: Audit trails are necessary to ensure that legal requirements such HIPAA and GDPR, which mandate the establishment of robust audit logging mechanisms to track access to patient data. Healthcare companies can help with regulatory compliance and show their dedication to patient privacy and security by keeping thorough audit trails.

MONITORING

Monitoring involves actively observing and analyzing system activities and events in real-time to identify anomalies, security threats, or policy violations. In the context of safeguarding patient privacy In the era of big data, monitoring is essential forproactively detecting and responding to potential security incidents or breaches that could compromise patient data.

a. Real-Time Detection: Monitoring systems continuously analyse system logs, network traffic, and user activities to detect unusual patterns or behaviours indicative of unauthorized access or suspicious activities. Real-time alerts and notifications enable security personnel to respond promptly to potential threats and mitigate risks before they escalate.

b. Threat Identification and Mitigation: Monitoring helps identify potential security threats, such as unauthorized access attempts, data exfiltration, or malware infections, that could compromise patient privacy. By monitoring network traffic, endpoint activities, and system logs, healthcare organizations can detect and respond to security incidents in a timely manner, minimizing the effect onpatient information and confidentiality.

c. Performance Optimization: In addition to security monitoring, healthcare organizations can leverage monitoring tools to optimize system performance, identify bottlenecks, and improve resource utilization. By monitoring system health and performance metrics, organizations can ensure the reliable and efficient operation of systems and applications while maintaining patient privacy and security.

By implementing robust audit trails and monitoring mechanisms, healthcare organizations can enhance patient privacy protections in the age of big data.These measures enable organizations to track access to patient data, detect unauthorized activities, and respond promptly to security incidents or

breaches, thereby safeguarding patient privacy and ensuring compliance with regulatory requirements. Additionally, audit trails and monitoring support ongoing security monitoring and performance optimization efforts, helping to maintain the integrity, availability, and confidentiality of patient data in healthcare environments.

4. PRIVACY-PRESERVING ANALYTICS

Privacy-preserving analytics refers to a set of techniques and methodologies designed to analyse data while simultaneously protecting the privacy and confidentiality of sensitive information contained within the data. This approach allows organizations to derive insights and value from data without exposing sensitive details that could compromise individuals' privacy. Privacy-preserving analytics is particularly relevant in contexts where data contains personally identifiable information (PII) or other sensitive attributes.

a. Data Aggregation and Masking*:* Privacy-preserving analytics techniques can involve aggregating patient data to produce statistical summaries or masked representations of the original data. Aggregated data can offer insightful information onpopulation health trends, treatment outcomes, and disease prevalence without exposing individual patient identities or sensitive details. Masking techniques, such as data anonymization or encryption, can further protect patient privacy by concealing identifiable information while retaining the utility of the data for analysis.

b. Secure Multiparty Computation (SMC): Secure multiparty computation allows multiple parties to jointly analyse data while keeping their individual inputs private. Without directly exchanging the underlying data, SMC facilitates collaborative analysis of patient data from various healthcare providers or research organizations in the context of healthcare. Each party can compute aggregate statistics or perform analytics on their respective datasets without revealing sensitive patient information to other parties, thereby safeguarding patient privacy.

c. Homomorphic Encryption: Homomorphic encryption is a cryptographic technique that allows computations to be performed without first decrypting the encrypted data first. In healthcare analytics, homomorphic encryption enables researchers or analysts to perform calculations on encrypted patient data while preserving confidentiality. This approach ensures that sensitive information remains protected throughout the analysis process, mitigating the risk of unauthorized access or disclosure.

d. Differential Privacy: Differential privacy techniques add noise to query results or statistical summaries to protect individual privacy while still allowing for meaningful analysis. In relation to medical treatment analytics, differential privacy can be applied to aggregate queries or machine learning techniques to make sure theinclusion or exclusion of a single patient's data does not significantly impact the overall results. This approach helps balance the need for data utility with the imperative to safeguard patient privacy in the age of big data era.

By leveraging privacy-preserving analytics techniques such as data aggregation, secure computation, encryption, and differential privacy, healthcare organizations and researchers can extract valuable insights from large datasets while minimizing privacy risks and maintaining compliance with regulatory requirements. These techniques enable the responsible application of big data in healthcare while upholding the highest standards of patient privacy and confidentiality.

5. DATA MINIMIZATION AND DE-IDENTIFICATION

Important privacy-enhancing strategies for safeguarding sensitive data include data minimization and de-identification, especially when working with huge datasets that include private or sensitive information. Each idea is explained as follows:

a. Data Minimization:

Data minimization involves limiting the collection, storage, and retention of patient data to only what is necessary for specific healthcare purposes. In the age of big data era, where vast amounts of patient data are generated and analysed, data minimization is essential for safeguarding patient privacy. Here's how data minimization applies:

b. Selective Data Collection: Healthcare organizations should only collect patient data that is directly relevant to providing care, managing treatment, or conducting research. By avoiding the assortment of unnecessary or excessive data, organizations can reduce the possibility ofprivacy breaches and unauthorized access to sensitive information.

c. Retention Limitation: Organizations should establish clear policies and procedures for the retention and disposal of patient data. Data should only be retained for as long as necessary to fulfil its intended purpose or comply with regulatory requirements. Retaining data beyond its useful lifespan increases the risk of privacy breaches and may expose patients to unnecessary privacy risks.

d. Data Masking and Aggregation: In cases where detailed patient data is not required for analysis, data minimization techniques such as aggregation and masking can be employed. Re-identification risk can be reduced without sacrificing the data's analytical value.and unauthorized access by aggregating the data into summary statistics or masking identifiers.

e. De-identification:
The process of eliminating or obscuring personally identifiable information (PII) from patient data in order to prevent individual identification is known as de-identification. De-identification techniques are commonly used to protect patient privacy while allowing for the analysis of anonymized healthcare datasets. Here's how de-identification applies:

f. Anonymization Techniques: Healthcare organizations can employ various anonymization techniques, such as generalization, suppression, and perturbation, to de-identify patient data. Generalization involves replacing specific values with broader categories to mask individual identities, while suppression removes certain identifiers entirely. Perturbation introduces random noise to data in order to stop there-identification of individuals.

g. Preserving Data Utility: De-identification aims to achieve equilibrium betweenprotecting patient privacy and preserving the utility of the data for analysis. By removing or masking identifiers while retaining the structure and integrity of the data, organizations can leverage de-identified datasets for research, analytics, and knowledge discovery without compromising patient privacy.

h. Compliance Considerations: De-identification is often a requirement for compliance with privacy regulations such as HIPAA and GDPR, which mandate the defence of patient privacy and confidentiality. By de-identifying patient data according toregulatory requirements, healthcare organizations can demonstrate their commitment to safeguarding patient privacy in the big data era.

In conclusion, in the age of big data age, data minimization and de-identification are critical tactics for protecting patient privacy. Healthcare organizations can reduce privacy risks while utilizing big data analytics to improve patient care, research, and outcomes by minimizing the collection and retention of patient data and using de-identification techniques to anonymize sensitive information.

6. ANONYMIZATION

By removing or changing personally identifiable information (PII) from datasets such that people cannot be re-identified, anonymization is a privacy-enhancing approach. Anonymization is a technique used to preserve privacy while preserving the ability to use data for analysis, research, and other acceptable uses. Anonymization can take many different forms, all of which aim to strike a compromise between the necessity of protecting individual privacy and the need for data value. Here are the main features and mechanisms of anonymization:

a. Protecting Patient Privacy: Anonymization helps safeguard patient privacy by ensuring that sensitive identifiers, such as names, addresses, and social security numbers, are removed or obfuscated from healthcare datasets. In the age of big data era, where vast amounts of patient data are collected and analysed, anonymization is necessary forpreventing unauthorized access to personal information and mitigating the possibility ofprivacy breaches.

b. Enabling Data Analysis: Despite the removal of identifiable information, anonymized datasets retain valuable useful informationfor research, analysis, and decision-making in healthcare. Anonymization allows researchers, clinicians, and data analysts to access and analyse healthcare data without compromising patient privacy. By anonymizing patient data, organizations can harness the power of big data analytics to identify trends, improve patient outcomes, and optimize healthcare delivery.

c. Balancing Privacy and Utility: Anonymization seeks to achieve equilibrium betweenprotecting patient privacy and preserving the utility of the data for analysis and research purposes. While anonymization techniques remove or mask identifiers, they aim to retain the integrity and structure of the data to ensure that it remains useful for statistical analysis, machine learning, and other analytical tasks. By anonymizing patient data, healthcare organizations can comply with privacy regulations, such as HIPAA and GDPR, while still deriving valuable insights from large datasets.

d. Compliance with Regulations: In order to comply with privacy legislation and standards governing the management of healthcare data, anonymization is frequently required. The preservation of patient privacy and the secure processing of personal health information are required by laws like the GDPR (General Data preservation Regulation) and HIPAA (Health Insurance Portability and Accountability Act). By anonymizing patient data, companies may comply with these legal obligations and maintain the highest standards of data security and privacy in the big data age.

CONCLUSION

In the age of big data in healthcare, protecting patient privacy is paramount for maintaining trust and adhering to ethical standards. While the digitization of healthcare data offers significant opportunities for improving patient care and advancing research, it's essential to balance these benefits with robust privacy measures. Through tactics such as encryption, anonymization, and data minimization, as well as stringent access controls and continuous monitoring, healthcare organizations can ensure patient data remains secure and confidential.

Healthcare businesses can reduce privacy risks and encourage responsible data stewardship by implementing a privacy-by-design strategy and including privacy safeguards into every phase of the data lifecycle. Enabling the revolutionary potential of big data analytics in healthcare through stakeholder collaboration, privacy-preserving technology implementation, and regulatory compliance further protects patient privacy.

REFERENCES

[1] Zhang and colleagues (2019). A framework for sharing healthcare data that protects privacy in the big data era. Big Data Transactions, IEEE, 5 (2). 10.1109/TBDATA.2018.2888558 is the doi.

[2] K. Abounelmehdi and colleagues (2018). Privacy rules and regulations for safeguarding patient privacy in the big data analytics era. Big Data Journal, 5(1). 10.1186/s40537-018-0132-6 is the DOI.

[3] Sweeney, L., and B. Malin (2012). Big data publishing for tailored medicine while protecting privacy. American Medical Informatics Association Journal, 19(3). 10.1136/amiajnl-2011-000425 is the doi

[4] Shafiq and colleagues (2019). Big data analytics that protects privacy: Opportunities and challenges. Computer Systems for Future Generation, 97. DOI: 10.1016/j.future.2019.01.044

[5] Dankar, F. K., and K. El Emam (2008). protecting patient privacy when exchanging health information. 9(6) Nature Reviews Genetics. 10.1038/nrg2386 [DOI] Zhou, X., & Coiera, E. (2015). Big data in health: challenges and opportunities. Big Data & Society, 2(2). DOI: 10.1177/2053951715612169

[6] Wang, Yu, & Associates (2018). Consortial blockchain enables safe and private data exchange for e-health systems. DOI: 10.1109/ACCESS.2018.2846699 IEEE Access, 6.

[7] Zhang, X., and Sun, J. (2016). protection of huge medical data privacy. Journal of Engineering in China, 2016 (8). 10.1155/2016/9260694 is the doi

[8] J. L. Fernández-Alemán et al. (2013). A comprehensive literature analysis on security and privacy in electronic health records. Biomedical Informatics Journal, 46(3). 10.1016/j.jbi.2012.12.003 is the doi

[9] Kayaalp, M. (2017). Privacy of patients in the big data era. Journal of Balkan Medicine, 34(1). 10.4274/balkanmedj.2017.0343 is the DOI.

[10] Coiera, E., and Zhou, X. (2015). Health and big data: possibilities and difficulties. Society & Big Data, 2 (2). Reference: 10.1177/2053951715612169