IJARIIT INTERNATIONAL JOURNAL OF ADVANCE RESEARCH, IDEAS AND INNOVATIONS IN TECHNOLOGY

ISSN: 2454-132X
Impact Factor: 6.078
(Volume 10, Issue 1 - V10I1-1197)
Available online at: https://www.ijariit.com

Cyber Analytics: Modelling the Factors Behind Healthcare Data Breaches for Smarter Security Solutions

Tosin Clement
clementtosin92@gmail.com
University of Louisville, Kentucky, USA

Callistus Obunadike
Callistusobunadike@gmail.com
Austin Peay State University, Clarksville, USA

Darlington C. Ekweli
darlingtonekweli@gmail.com
University of the Potomac, Washington DC USA

Oluomachi E. Ejiofor
pearloluomachi@gmail.com
Austin Peay State University, Clarksville, USA

Oluwadamilola Ogunleye
ogunleyo23@gwu.edu
George Washington University, Washington DC, USA

Simo Sevidzem Yufenyuy
syufenyuy@my.apsu.edu
Austin Peay State University, Clarksville, USA

Chukwu I. Nnaji
chukwunnaji@gmail.com
Austin Peay State University, Clarksville, USA

Chinenye J. Obunadike
josephcobunadike@gmail.com
Anambra State University, Uli Nigeria

ABSTRACT

This study employs a comprehensive methodology to analyze healthcare data breaches in the United States, utilizing information extracted from the U.S. Department of Health and Human Services Portal. The unbalanced nature of the data across different years is addressed through meticulous examination of breach occurrences, encompassing diverse factors such as state, covered entity type, affected individuals, breach type, and entity classification. The results section unveils key insights into the prevalence and impact of healthcare data breaches. Hacking and IT incidents emerge as the predominant breach type, significantly affecting individuals, followed closely by unauthorized access/disclosure and theft. The study further dissects the data by business type, revealing that business associates and healthcare providers bear the brunt of breaches, with health plans and healthcare clearing houses also facing substantial challenges. The study conducted cyber analytics on the factor behind healthcare data breaches for smarter security solutions. This is based on the backdrop of increasing cybercrime in the United States. The study utilized secondary data, which includes indicators such type of breach, location of breach, the number of individuals affected, business type, and the time of cyberattack. The findings revealed that hacking and information technology incidents are the most prevailing cyberattack on healthcare data, with healthcare providers and business associate being the most affected entity. The findings also revealed that network server and email are the major location of healthcare data breached. Furthermore, the data indicated that there is more

breach in 2023 than other years, indicating a significant rise in cyberattacks in the healthcare. It was suggested that healthcare entities need to develop and regularly update incident response plans to ensure a swift and effective response in the event of a cybersecurity breach, which should include clear communication strategies to prevent losing data to cybercriminals. The concentration of breaches in specific entities, states, and quarters underscores the diverse and pervasive nature of cybersecurity challenges in the healthcare sector. Continuous efforts to enhance cybersecurity frameworks are deemed critical to safeguard sensitive healthcare data and protect individuals' privacy.

KEYWORDS: *Cyber Analytics, Cyber Security, Healthcare Analytics, Data Breach, Security Solution*

I. INTRODUCTION

The advent of digital technologies has brought about a revolutionary period in healthcare, fundamentally altering the conventional framework of patient care and information administration (Mcleud & Dolezel, 2018). The growing dependence on digital technologies represents a significant change towards healthcare systems that are more effective, networked, and driven by data. Electronic Health Records (EHRs) are a fundamental aspect of the digital revolution, simplifying the storing and retrieval of patient data (Seh, et al., 2020). Healthcare practitioners now have convenient access to extensive medical records, treatment strategies, and diagnostic information, enabling them to make better-informed decisions and deliver individualized care to patients (Oloyede, 2023).

The use of telehealth and telemedicine has become a crucial factor, particularly in recent times, highlighting the flexibility and availability of healthcare services (Li, Xiao, & Zhang, 2023). Telehealth utilizes digital platforms to offer remote consultations, monitor patient vital signs, and facilitate prompt interventions, overcoming geographical limitations and improving healthcare accessibility for various groups (Dart & Ahmed, 2023). The emergence of wearable gadgets and Internet of Things (IoT) technology intensifies this dependence on digital solutions. Wearables offer immediate health data, enabling users to actively engage in their well-being by tracking metrics such as physical activity, heart rate, and sleep habits. The interconnectedness of these devices creates a holistic digital ecosystem that not only encourages proactive healthcare but also enables timely intervention and management of chronic diseases (Abouelmehdi et al., 2018).

The significance of healthcare data in the contemporary medical field cannot be exaggerated. Healthcare decision-making relies heavily on patient health records, treatment histories, diagnostic pictures, and a wealth of sensitive information (Oloyede, et al., 2023). The extensive collection of data not only provides information for individual patient treatment but also adds to wider public health efforts, medical research, and healthcare system administration. Nevertheless, the abundance of information in the healthcare sector renders it an appealing target for cyber assaults (Reddy et al., 2021). The vulnerabilities related to healthcare data are complex, involving both the technological and human components of data management. Cybercriminals, who are aware of the significant worth of health records, utilize several strategies like ransomware attacks, phishing scams, and data breaches to take advantage of these weaknesses (Javoid et al., 2023).

The ramifications of a successful cyberattack on healthcare data are significant and go beyond mere monetary damages. The compromise of patient privacy undermines confidence in the healthcare system, while the possibility of medical identity theft presents significant dangers to individuals (Sardi et al, 2020). Furthermore, the interconnectivity of healthcare systems and the growing dependence on Internet of Things (IoT) devices also broaden the potential targets for cyber-attacks. These technologies, which include medical sensors and connected medical equipment, create more opportunities for cyber risks to occur (Raghupathi et al., 2023). The interaction between technology, confidential health data, and the complex network of healthcare systems forms a dynamic environment in which cybersecurity is not only a technological requirement but also a core ethical responsibility (Li *et al*., 2023). With the increasing digitization of healthcare, it is crucial to continuously prioritize the protection of patient data by implementing strong cybersecurity frameworks and actively resolving weaknesses in this essential area (Almulihi, et al., 2022).

The increasing worry of healthcare data breaches has become a central issue in the United States, highlighting the greater difficulties encountered by the healthcare industry in safeguarding sensitive patient information (Ignatowski, 2021). HIPAA is a regulatory framework that aims to safeguard the confidentiality and integrity of health information. Nevertheless, despite the implementation of these protective measures, the healthcare industry in the United States has witnessed a significant increase in incidents of unauthorized access to sensitive information, leading to heightened endeavors to rectify existing weaknesses (Oloyede, 2023).

An important part of the concern is the magnitude and frequency of reported breaches. The vulnerability of the healthcare sector to cyber threats has been highlighted by notable instances affecting prominent healthcare organizations, insurance providers, and medical institutes (Dean, 2023). The stolen data frequently encompasses not just personal health information but also financial particulars, resulting in a complex risk environment. The financial consequences for impacted persons in the United States are significant (Mcleud & Dolezel, 2018). Medical identity theft can lead to victims experiencing financial difficulties due to fake insurance claims, illegal medical treatments, and prescription fraud. Furthermore, the possibility of sensitive health information being revealed raises worries regarding its potential misuse in different scenarios, such as employment determinations and insurance benefits (Abouelmehdi *et al.*, 2018).

The structure of the U.S. healthcare system, characterized by extensive networks of healthcare providers, insurers, and third-party service providers, adds to the intricacy of cybersecurity issues (Li *et al.*, 2023). The growth of telehealth services, electronic record systems, and the increasing number of Internet of Things (IoT) devices make it easier for cyber-attacks to occur. This means that a thorough and flexible cybersecurity strategy is necessary. Ransomware attacks, a widespread type of cyber menace, have specifically targeted healthcare organizations in the United States, resulting in interruptions to vital medical services and prompting concerns about the sector's ability to withstand such attacks (Almulihi, et al., 2022). The possibility of extensive data breaches compromising patient care, disrupting healthcare operations, and undermining public confidence in the healthcare system is an urgent worry that requires focused efforts to strengthen cybersecurity defenses (Seh, et al., 2020).

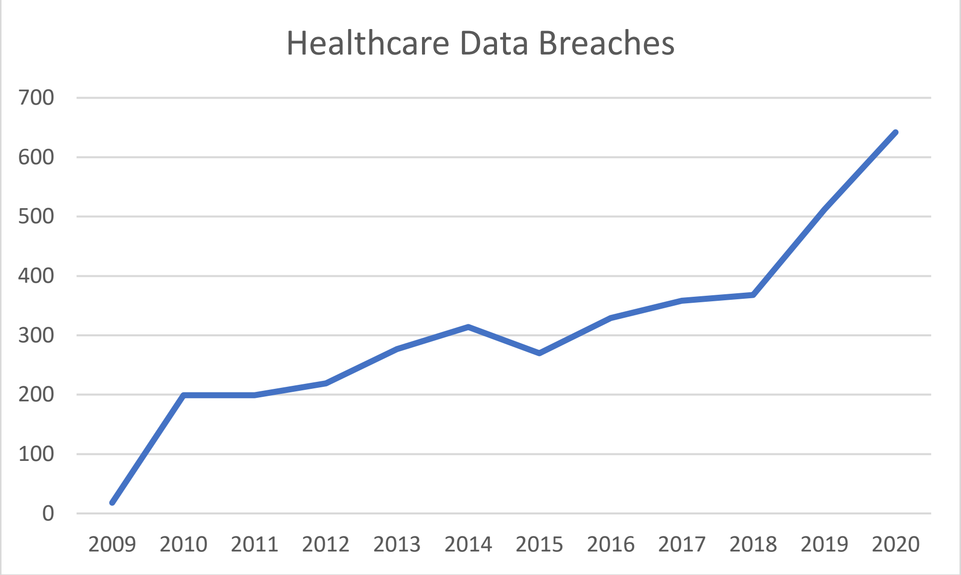

Figure 1: Healthcare Data Breaches 2009-2020

People are at risk of having their privacy violated as personal health records become accessible to unauthorized parties. The erosion of privacy not only puts the confidentiality of medical information at risk, but also creates emotions of vulnerability and dread among patients (Ignatowski, 2021). The stolen medical identities result in significant financial consequences, which exacerbate the effects by giving rise to fraudulent actions, including unapproved medical operations and prescription theft (Dean, 2023). The financial load presents difficulties for persons impacted by identity theft, as they must negotiate the consequences of identity misuse and work towards resolving bogus claims to regain their financial stability. Furthermore, the decline in confidence in healthcare institutions has resulted in patients becoming cautious about disclosing sensitive information, which could hinder correct diagnoses and successful treatment (Raghupathi *et al.*, 2023).

Compromised patient information can cause significant disruptions to healthcare institutions in the U.S., especially when targeted by ransomware attacks that have the potential to incapacitate essential medical services (Sardi *et al.*, 2020). These disruptions not only affect the provision of medical treatment and emergency services, but also put pressure on the general operation of healthcare facilities. Healthcare organizations face an additional level of complexity due to regulatory scrutiny and legal ramifications (Ayereby, 2018). They may be subject to fines and penalties for violating data protection requirements. Data breaches not only compromise the security of healthcare organizations but also undermine their reputation within the sector. That makes it essential to adopt a comprehensive strategy that includes strong cybersecurity protocols, strict adherence to legal requirements, and initiatives aimed at restoring confidence between healthcare practitioners and their patients (Almalawi *et al.*, 2023).

The need for improved cybersecurity measures in the health sector is crucial because of the increasing frequency and complexity of cyber assaults that specifically target sensitive patient information (Javoid *et al.*, 2023). The susceptibility to cyberattacks increases dramatically as healthcare firms digitize and integrate their systems. An effective cybersecurity strategy is crucial for protecting patient privacy, ensuring uninterrupted operations, and maintaining the credibility of healthcare organizations in the midst of ever-changing and persistent cyber threats (Li *et al.*, 2023). The research goal is to utilize cyber analytics to model the elements that contribute to healthcare data breaches. The study seeks to utilize sophisticated analytical methods to uncover crucial variables and trends that are responsible for security vulnerabilities in the healthcare industry. This will enable the creation of more intelligent security solutions that can successfully mitigate and prevent data breaches.

II. LITERATURE REVIEW

Cybersecurity involves safeguarding computer systems, networks, and data from unauthorized access, attacks, and harm to guarantee that confidentiality, integrity, and availability are maintained (Oloyede, et al., 2023). Cybersecurity, within the realm of cyber analytics, entails utilizing sophisticated analytical methodologies and tools to scrutinize extensive volumes of data, discern trends, and ascertain potential threats or weaknesses. Cyber analytics is essential for improving cybersecurity by offering practical insights, anticipating future threats, and enabling proactive measures to efficiently avoid and address security incidents (Dart & Ahmed, 2023).

Cyber Analytics

Cyber analytics is a field within cybersecurity that utilizes complex analytical methods to examine large amounts of data produced in digital environments (Mcleud & Dolezel, 2018). Cyber analytics use data-driven techniques, machine learning algorithms, and behavioral analysis to detect trends, abnormalities, and potential security concerns in intricate networks and systems. Cyber analytics examines both historical and real-time data to identify not only familiar cyber dangers but also emerging and changing methods of assault (Javoid, Haleem, Singh, & Suman, 2023). It has a crucial role in improving proactive cybersecurity measures, providing firms with the ability to anticipate, stop, and promptly address cyber incidents. Cyber analytics offers valuable insights by consistently monitoring and analyzing data, enabling security experts to comprehend the dynamic nature of potential threats (Reddy, Elsayed, Elsayed, & Oser, 2021). This proactive strategy enables firms to strengthen their security measures, adjust to evolving threats, and enhance their overall ability to withstand cyber-attacks in a world that is becoming more linked and reliant on digital technology.

Overview of Healthcare Data Breaches

Healthcare data breaches present substantial obstacles to maintaining the secrecy, accuracy, and accessibility of delicate patient information in the healthcare industry. Due to the growing digitization of health records and the widespread adoption of electronic health information systems, the amount of data at risk of being breached has increased significantly (Cremer, et al., 2022). The ramifications of these breaches go beyond the violation of personal privacy and have the potential to endanger patient safety and undermine public confidence in healthcare organizations (Ayereby, 2018).

A significant element that contributes to healthcare data breaches is the high value of medical records on the black market, where stolen healthcare data can command greater prices compared to other types of personal information (Javoid *et al.*, 2023). Cybercriminals focus on healthcare businesses to take advantage of the extensive amount of information included in medical records, such as personally identifiable information (PII), financial data, and insurance details (Dean, 2023). Moreover, the interlinked structure of healthcare systems expands the potential for attacks, enabling skilled malicious actors to take advantage of weaknesses in networks, medical equipment, and third-party service providers. Healthcare data breaches can have significant consequences, including identity theft, fraudulent behavior, and potential injury to patients if their medical records are tampered with or exploited (Ismagilova *et al.*, 2022).

In addition, the healthcare industry is subject to a regulatory framework that includes standards like the Health Insurance Portability and Accountability Act (HIPAA), which requires strict security measures to safeguard patient data (Almulihi, et al., 2022). Failure to comply with these requirements not only subjects' organizations to legal consequences but also erodes the confidence patients have in healthcare professionals to protect their confidential data. To effectively tackle the increasing number of healthcare data breaches, a comprehensive strategy is needed. This strategy should include advanced cybersecurity measures, strong adherence to regulations, and continuous education and training for healthcare professionals to effectively deal with the evolving threat landscape (Abouelmehdi *et al.*, 2018).

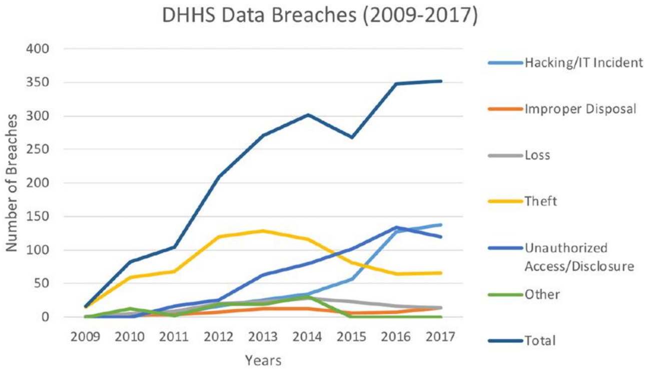

Figure 2: DHHS Data Breaches 2009-2017

Cyber-Analytics in Healthcare

Cyber analytics is crucial in strengthening the cybersecurity position of healthcare companies by providing proactive measures to detect, prevent, and address emerging cyber threats (Ismagilova *et al.*, 2022). Given the rapid increase in healthcare data and the growing complexity of digital healthcare systems, conventional security methods are inadequate in safeguarding against advanced assaults. Cyber analytics use cutting-edge technology like artificial intelligence, machine learning, and behavioral analytics to evaluate large amounts of data in real-time (Raghupathi *et al.*, 2023). It identifies irregularities and patterns that may indicate possible security risks. By adopting this proactive approach, healthcare companies may surpass reactive responses and effectively anticipate and mitigate hazards before they worsen.

Cyber analytics plays a vital role in the healthcare industry by securing patient data, maintaining the confidentiality of medical records, and defending against ransomware attacks and other cyber threats that have the potential to interrupt essential healthcare services (Karunarathne *et al.*, 2021). Cyber analytics enables healthcare workers to enhance security measures and promptly address emerging threats by consistently monitoring network activities, user behavior, and system vulnerabilities, and providing actionable insights (Cremer, et al., 2022). The utilization of cyber analytics enables adherence to regulatory mandates, such as the Health Insurance Portability and Accountability Act (HIPAA), by guaranteeing strong measures to protect confidential patient data. With the growing adoption of digital transformation in healthcare businesses, it is crucial to use cyber analytics to effectively manage the intricate realm of cyber threats and safeguard the trust and welfare of patients (Ignatowski, 2021).

Factors Contributing to Healthcare Data Breaches

The occurrence of healthcare data breaches can be attributed to a confluence of variables, which are indicative of the distinctive difficulties encountered by the healthcare sector in protecting confidential patient data (Almalawi *et al.*, 2023).

a) **High-Value Target:** The healthcare industry possesses a wealth of valuable information, rendering it an appealing objective for cybercriminals. Patient records encompass a vast array of information, encompassing personally identifiable information (PII), medical histories, and financial particulars (Ayereby, 2018). As a result, healthcare organizations become attractive targets for illicit activities in the underground market.

b) **Complex Ecosystem:** The interconnections of healthcare systems and the widespread utilization of electronic health data give rise to a complex ecosystem that harbors multiple vulnerable points susceptible to breaches. The wide-ranging network, encompassing hospitals, clinics, laboratories, and third-party service providers, amplifies the potential targets for cybercriminals (Sardi *et al.*, 2020).

c) **Obsolete Infrastructure:** Numerous healthcare facilities continue to depend on legacy systems and antiquated infrastructure that may lack comprehensive cybersecurity protections. Outdated software and hardware may have unaddressed vulnerabilities, which can be exploited by attackers to take advantage of deficiencies (Ayereby, 2018).

d) **Human Error and Insider Threats:** Human error and insider threats can also lead to healthcare data breaches. These breaches may occur due to mistakes in configuring security settings, mishandling of devices, or unintentional disclosure of sensitive information (Dart & Ahmed, 2023). Furthermore, the presence of insider threats, whether deliberate or accidental, poses a significant danger as employees or associated individuals undermine the security of data.

e) **Ransomware Attacks:** The increasing prevalence of ransomware attacks in the healthcare industry is a major cause for concern. Cybercriminals utilize malicious software to encrypt valuable data, and then demand a ransom in exchange for its decryption (Li *et al.*, 2023). Such disruptions can cause significant disruptions to healthcare operations, damage the quality of patient treatment, and result in the unauthorized extraction of data.

f) **Lack of Cybersecurity Awareness:** Cybersecurity deficiency Cybersecurity vulnerabilities in the healthcare sector are exacerbated by a lack of proper awareness and training among healthcare practitioners. Personnel may become targets of phishing attempts or unintentionally participate in activities that jeopardize cybersecurity (Dart & Ahmed, 2023).

The Role of Regulatory and Compliance Bodies

Regulatory and compliance entities have a vital function in determining and implementing cybersecurity regulations in the healthcare sector. Their supervision is crucial in guaranteeing that healthcare firms comply with rigorous requirements aimed at safeguarding patient data and upholding the integrity of healthcare systems (Ismagilova *et al.*, 2022). The Health Insurance Portability and Accountability Act (HIPAA) is a significant regulatory framework in the United States. HIPAA sets forth nationwide regulations to secure patient health information (PHI) and requires measures to prevent illegal access, disclosure, and utilization of this information (Oloyede, et al., 2023). Healthcare providers, health plans, and healthcare clearinghouses, known as covered entities, are required to adhere to HIPAA standards in order to protect patient privacy and ensure security.

The Office for Civil Rights (OCR) is an important regulatory agency responsible for enforcing compliance with the Health Insurance Portability and Accountability Act (HIPAA). The OCR is responsible for examining instances of unauthorized access to data, performing thorough examinations of records, and enforcing consequences for failure to adhere to regulations (Almulihi, et al., 2022). The regulatory framework grants people authority over their health information and establishes the foundation for a secure and private healthcare setting. Healthcare organizations are required to comply with both federal regulations and state-specific regulations, as well as adhere to international standards (Reddy *et al.*, 2021). The General Data Protection Regulation (GDPR) has a significant influence on healthcare enterprises that operate on a global scale. The General Data Protection Regulation (GDPR) establishes stringent guidelines for the handling and safeguarding of personal data, which encompasses sensitive health-related information.

The regulatory and compliance agencies not only establish norms, but also actively participate in monitoring and enforcing them. Failure to adhere to regulations can lead to significant repercussions, such as monetary penalties, legal proceedings, harm to reputation, and erosion of trust among patients and stakeholders (Karunarathne *et al.*, 2021). The regulatory environment is always changing to deal with new cybersecurity issues, highlighting the need for a proactive and flexible approach to compliance. It is crucial for regulatory bodies, healthcare organizations, and cybersecurity experts to collaborate to anticipate and address emerging threats and to consistently enhance security measures in the ever-changing digital healthcare environment (Li *et al.*, 2023).

Challenges of Application of Cyber Analytics in the Healthcare Industry

Healthcare data is extremely diverse and complicated, which is a big challenge. Many different types of data are created by healthcare systems. This includes EHRs, medical imaging files, billing information, and a lot more besides (Cheng *et al.*, 2017). Conventional analytics methods face a formidable obstacle when trying to integrate and analyze such a heterogeneous data landscape. The smooth implementation of advanced analytics can be impeded by problems with data format and system interoperability.

The necessity of healthcare-related real-time analysis and decision-making presents still another obstacle. Cyber analytics can throw light on important issues, but healthcare organizations frequently need quick answers to new dangers or outliers (Cremer, et al., 2022). The security of patients and their data can be jeopardized if analytics procedures are slow or ineffective. Healthcare businesses must build analytics solutions that work in real-time to proactively handle security issues and mitigate potential risks in response to the requirement for quick identification and reaction to cyber threats (Ignatowski, 2021).

Another major obstacle to successful cyber analytics adoption in healthcare is the scarcity of qualified cybersecurity experts. Cybersecurity skills are scarce, which is preventing the use of advanced analytics tools and approaches (Almalawi *et al.*, 2023). It could be difficult for businesses to recruit and keep employees with the right kind of knowledge to handle cyber analytics and make sense of the results. The healthcare industry's cybersecurity landscape is complicated, and filling this gap will necessitate concerted training and education initiatives.

Tackling these difficulties calls for a comprehensive strategy that merges advancements in technology, training of healthcare workers, and cooperation between cybersecurity and healthcare organizations. To fully realize cyber analytics', promise of strengthening the healthcare industry's security posture, several obstacles must be overcome (Mcleud & Dolezel, 2018).

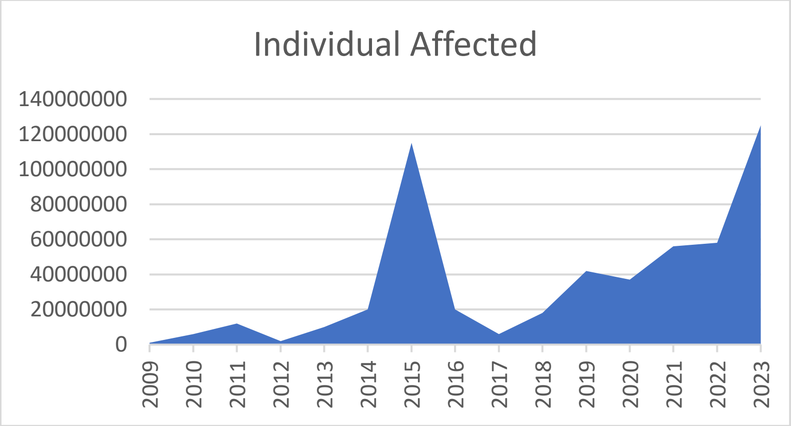

Figure 3: Individual Affected 2009 - 2023

Ransomware Attacks in the Healthcare

a) RYUK: A cybercrime ring that infiltrated 400+ US healthcare facilities using RYUK ransomware, which mainly targets financial advantages. Significantly, it was the third most common type of ransomware assault in 2020, with victims allegedly spending almost $61 million to decrypt their data (Dean, 2023). After systematically trying to erase files and backup data, the ransomware copies these files in shadow, disrupts security services, disables Windows Automatic Startup Repair, changes boot status settings, and leaves a ransom notice on the victim's screen (Cheng *et al.*, 2017).

   The ransom demands a specific amount to decrypt the victim's data; the message provides contact information for the ransom buyer. The WannaCry virus crisis of 2017 was similar in nature, and it impacted almost 40% of healthcare companies (Karunarathne *et al.*, 2021).

b) Trickbot: Trickbot, UNC1878, or Team9, was created by the Wizard Spider cybercrime group. U.S. hospitals, clinics, and government agencies are the deliberate targets of this gang. As a botnet that provides backdoor access to compromised machines (Sardi *et al.*, 2020), Trickbot uses a technology called anchor-DNS. The infected system can then be used to infect additional computers using this access. The four main ways that Trickbot injects malware are through spear phishing, secondary payloads, vulnerabilities in networks, and malvertising. The importance of cybersecurity measures in securing sensitive information is shown by the group's focus on crucial sectors (Raghupathi *et al.* , 2023).

Theoretical Review

The study employed the Routine Activity Theory (RAT) to explain using cyber analytics to model the factors behind healthcare data breaches.

Routine Activity Theory

This theory was developed by Cohen and Felson (1979). It is a criminological theory that focuses on everyday activities of individuals and how they create opportunities for criminal behavior. Routine Activity theory suggests that crime occurs when three elements converge in time and space: a motivated offender, a suitable target, and the absence of a capable guardian. When it comes to healthcare data breaches (Wilcox, 2015). The theory can provide light on the everyday tasks that could leave the healthcare industry vulnerable to cyberattacks. An effective cybersecurity system acts as a capable guardian, a motivated offender as a hostile actor aiming to get unauthorized access to private healthcare information, and a suitable target as a susceptible data infrastructure (Argun & Daglar, 2016). Applying routine activity theory allows you to examine healthcare systems for trends and routines that could leave them vulnerable to data breaches. Since this theory places an emphasis on comprehending the dynamics of criminal chances in a certain context, it is in line with the idea of cyber analytics.

Empirical Studies

Raghupathi et al. (2023) examine the occurrence of data breaches in healthcare provider settings regarding patient data. The study analyzes the distribution of breaches across states, explores main causes and types of breaches, and assesses their impact on providers and patients. Findings highlight diverse data breach victims, with network servers being the primary location for common breaches. Recommendations for proactive measures, including regulatory compliance and network server monitoring, are provided.

Oloyede, et al. (2023) examines the relationship between cybersecurity and identity theft in the United States from 2001 to 2021. Trend analysis and Chi-Square tests were employed using time series data from various sources. The findings indicate a consistent rise in consumer complaints about identity theft and increasing spending on cybersecurity. However, the Chi-Square analysis suggests that cybersecurity does not have a significant impact on identity theft. The study recommends heightened public awareness, increased organizational investment in security systems, and international collaboration to combat fraudsters.

Reddy et al, (2021) highlighted the critical issue of healthcare data breaches, particularly through ransomware incidents, posing a significant threat to patient privacy despite established security standards like HIPAA. The study focuses on understanding the factors leading to such breaches and their impact on patients and healthcare providers. Additionally, it reviews current solutions aimed at enhancing healthcare security and evaluates their effectiveness. The analysis centers on notable ransomware attacks in the U.S. from 2015 to 2020, utilizing data from diverse academic and business sources to investigate the root causes of healthcare data breaches.

Seh, et al. (2020) explored categories of healthcare data breaches, emphasizing hacking/IT incidents and unauthorized internal disclosures. The analysis indicates a rising trend in breach frequency, exposed records, and financial losses. Valuable healthcare data attracts misuse, prompting the study to employ time series analysis methods for forecasting trends and costs associated with data breaches, with the simple moving average method proving more reliable. It was concluded that healthcare industry is a primary target of numerous cyber threats.

Mcleud & Dolezel (2018) construct a model identifying factors related to healthcare data breaches. Variables, such as healthcare facilities' exposure and security levels, along with organizational factors, were operationalized. The outcome variable was binary, representing the presence or absence of a data breach. Given the risks associated with healthcare data breaches, including personal health information exposure, the study holds significance in the healthcare field. Data from the Department of Health and Human Services and a national database were analyzed using binary logistic regression, revealing several exposure, security, and organizational factors significantly linked to healthcare data breaches.

III. METHODOLOGY

The study extract healthcare data breaches data from the U.S. Department of Health and Human Services Portal cover occurrences between 2021 and 2023. The breach reports results feature information on the healthcare firms, location, individuals affected, type of breach and where the breach occurred. The occurrences varied from days, months, and years, indicating that the data is unbalanced across the years. This database only includes data breaches that started in the United States. While this could be a drawback, the data set is comprehensive enough to allow for conclusions to be applied to a broader context. The study examined every record in the database, considering factors including state, kind of covered entity, affected persons, type of breach, and entity type. Steps in the process include gathering data, choosing variables, presenting data, selecting an analytics platform, tools, and finally, implementing analytics. We used the.xlsx format to retrieve the raw data.

IV. RESULTS

This section presents the results from the data analyzed. Table 1 presents the most common breach type that individuals suffered in the United States. Hacking and IT incident is the prominent cyber-attack on healthcare data, which was followed by unauthorized access or disclose of vital information.

Table 1: Type of Breach and Individuals Affected

Type of Breach	*Individuals Affected*	*Frequency*
Hacking/IT Incident	133366056	700
Improper Disposal	5546	3
Loss	1745	2
Theft	55147	11
Unauthorized Access/Disclosure	8429014	93
Grand Total	141857508	809

Theft is among the common healthcare data breach across the United States. However, it is important to mention other types of breach including improper disposal and loss, though they are relatively small when compared to hacking and IT incidents across healthcare organizations.

Table 2: Business Type and Individuals Affected

Row Labels	*Individuals Affected*	*Frequency*
Business Associate	73567896	150
Health Plan	16862776	116
Healthcare Clearing House	6242	4
Healthcare Provider	51410594	538
Grand Total	141847508	808

Table 2 presents distribution of entity with respect to number of individuals affected by data breach. It can be seen from the table that more individuals are affected within the business associated entity type. The outcome revealed that healthcare providers are also largely affected by data breaches, which has led to loss of data of millions of individuals. It was also observed that health plan is another entity with prominent cases of data breaches in the United States. On the other hand, healthcare clearing house is the least affected.

Table 3: Location of Breached Information

Location of Breached Information	*Individuals Affected*	*Frequency*
Network Server	138244482	582
Email	3107351	172
Electronic Medical Record	256328	26
Paper/Films	140416	17
Laptop	55733	7
Desktop Computer	53198	5
Grand Total	141857508	809

Table 3 presents information on the location of breached information and the corresponding number of individuals affected, along with the frequency of breaches in each category. The Network Server is the most significant source of breached information, impacting a substantial number of individuals, with a frequency of 582 incidents. Email breaches affected a significant number of individuals, breaches related to electronic medical records impacted a significant number of people, while incident involving paper and films occurred 17 times affected hundreds of individuals. Breached information relating to laptop and desktop are relatively small with occurrence of 7 and 5 respectively.

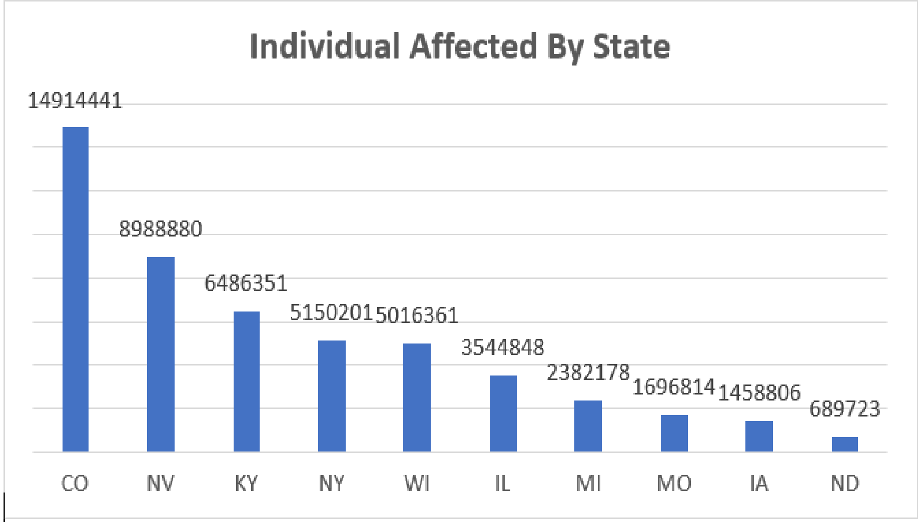

Figure 4: Number of Individuals Affected by State

Figure 4 presents the top ten (10) states affected by healthcare data breach in the United States. As presented in Figure 4, Colorado tops the list with a substantial number of individuals affected, indicating a significant healthcare data breach impact in the state. This was followed by Nevada, which suggests that there is high occurrence of healthcare data breach with the state. The third most affected state is Kentucky, with a substantial number of people affected, signifying a considerable impact on the privacy and security of healthcare data in Kentucky. New York is placed as the fourth most affected state with about 5.2 million individuals affected by cyberattack. Other states include Wisconsin, Illinois, Michigan, Iowa, and North Dakota respectively.

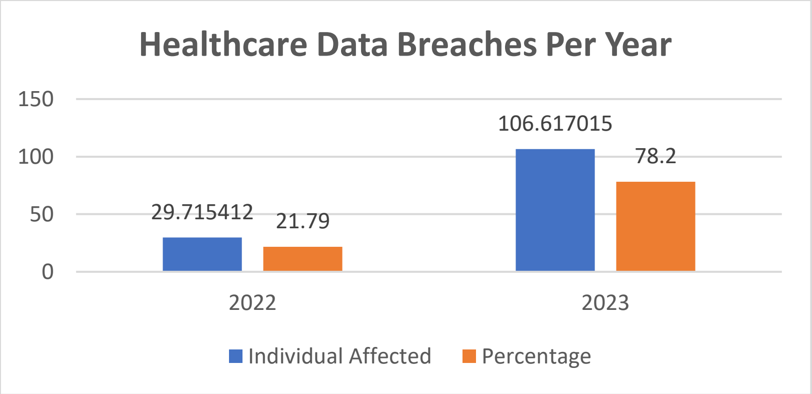

Figure 5: Healthcare Data Breaches for 2022 and 2023

As presented in Figure 5, incidence of data breaches was high in 2023, with about 106.61 billion individuals affected compared to 29.72 billion individuals affected in 2022. The implication of this is that the healthcare niche was highly targeted by cyber criminals in 2023.

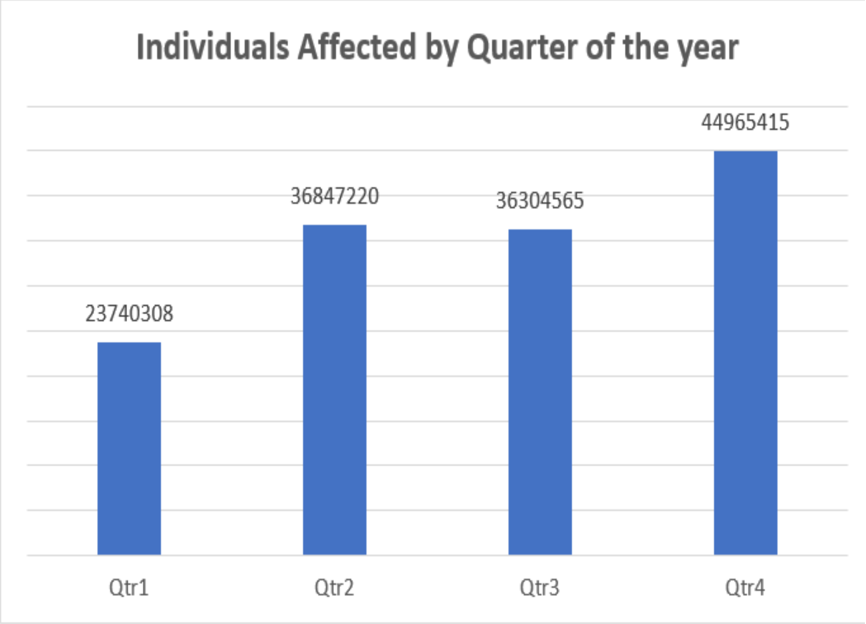

Figure 6: Number of Individuals Affected by Quarter of the Year

As presented in Figure 6, many healthcare entities experienced data breaches in the last quarter of the years, this was followed the second quarter of the year, which recorded 3.68 billion individuals affected by data breaches and the third quarter with 3.63 billion individuals affected. The first quarter of the year had the minimum number of individuals affected in terms of data breaches. The figures indicate fluctuating patterns throughout the year. While Qtr1 and Qtr3 show relatively lower sums, Qtr2 demonstrates a notable increase. However, the highest sum is observed in Qtr4, suggesting a potential seasonality or periodicity in healthcare data breach occurrences. This quarterly analysis underscores the importance of continuous vigilance and proactive measures to address and mitigate healthcare data breaches, with a particular emphasis on periods of heightened vulnerability.

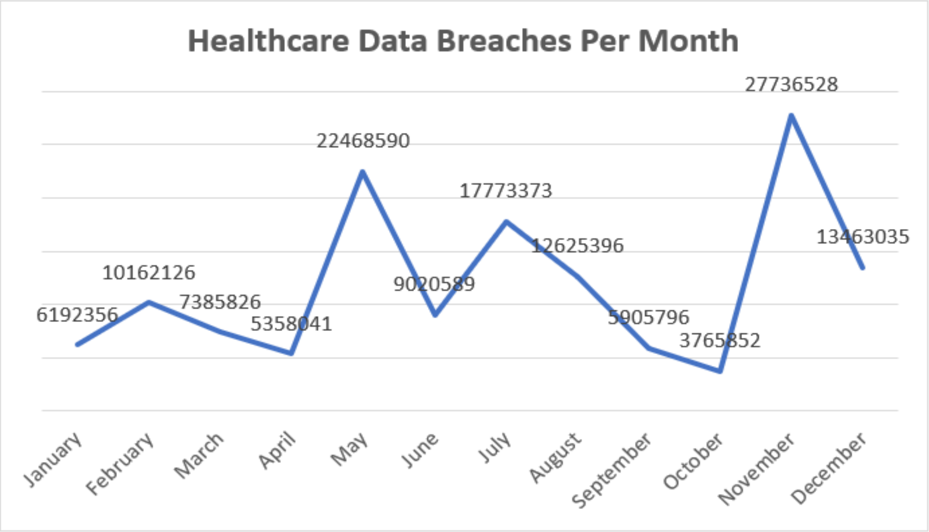

Figure 7: Healthcare Data Breaches Per Month

The data depicts the sum of individuals affected by healthcare data breaches across different months. The figures exhibit variations in the number of individuals affected, reflecting potential patterns or trends in healthcare data breach occurrences. Notably, certain months, such as May, June, and November, show significantly higher sums, indicating potential periods of increased vulnerability or targeted activities. Conversely, other months, such as April and October, demonstrate lower figures. This monthly analysis emphasizes the need for tailored cybersecurity measures that consider temporal patterns, allowing organizations to strengthen their defenses during periods of heightened risk. It underscores the dynamic nature of healthcare data breaches and the importance of adaptive strategies to safeguard individuals' sensitive information throughout the year.

V. DISCUSSION AND CONCLUSION

The findings revealed that has been consist rise in healthcare data breaches in the United States, with a significant difference in occurrence between 2022 and 2023. However, there are varying insights obtained from the findings. It was indicated that hacking and IT incidences are the most reoccurring breach that occurred, with unauthorized disclosure being another prominent data breach. It could be seen that identity theft is on the rise, which may call for stringent measures before it escalates. Business Associates and Healthcare Providers emerge as significant contributors to the overall count, collectively affecting a substantial number of individuals. This gives an insight into the entities that cyber perpetuators are targeting. Organizations within this niche may need to strengthen their security. It turns out that most breaches, which impact a lot of people, originate on the network server. Although they account for a smaller percentage than breaches involving network servers, intrusions involving email nevertheless add to the total effect. In terms of the states affected, Colorado and Nevada top the state with the most affected individual. This could be because of the concentration of healthcare organizations in these regions. The findings revealed that the fourth quarter experienced the highest occurrences of healthcare data breaches, and further analysis indicated that November is the month with the highest figure. This suggests that data breaches occur around the end of the year towards the festive period. The findings underscore the pervasive and diverse nature of cybersecurity challenges in the healthcare sector. The substantial number of individuals affected across different breach sources, including network servers, email, and electronic medical records, highlights the urgency of robust cybersecurity measures. The concentration of breaches in certain states and quarters emphasizes the need for targeted interventions and heightened vigilance during specific periods. The prevalence of breaches across various locations, such as laptops and desktop computers, underscores the necessity for comprehensive cybersecurity strategies to address vulnerabilities in different information channels. These implications emphasize the critical importance of continuous efforts to enhance cybersecurity frameworks and protect individuals' sensitive healthcare data.

Recommendations

Given the insights from the analysis, it is imperative to strengthen network security. This implies that healthcare organization should make it a top priority, which can be done through implementation of advanced encryption methods, regularly updating security protocols, and conducting thorough vulnerability assessments to identify, and address potential weak points in the network infrastructure. Human error is a significant factor in cybersecurity breaches. Healthcare organizations must establish regular and comprehensive training programs for employees to raise awareness about phishing attacks, social engineering tactics, and best practices for safeguarding sensitive information. Well-informed staff can act as a crucial line of defense against potential threats. Lastly, healthcare entities need to develop and regularly update incident response plans to ensure a swift and effective response in the event of a cybersecurity breach. This should include clear communication strategies, coordination with relevant authorities, and steps for mitigating the impact on affected individuals. Regularly conduct drills to test the efficiency of these plans and make necessary adjustments based on the outcomes.

VI. REFERENCES

Abouelmehdi, K., Beni-Hessane, A., & Khaloufi, H. (2018). Big healthcare data: preserving security and privacy. *Journal of Big Data, 5*, 1-8.

Almalawi, A., Khan, A. I., Alsolami, F., Abushark, Y. B., & Alfakeeh, A. S. (2023). Managing Security of Healthcare Data for a Modern Healthcare System. *Journal of Sensors, 23*(7), 3612.

Almulihi, A. H., Alassery, F., Khan, A. I., Shukla, S., Gupta, B. K., & Kumar, R. (2022). Analyzing the Implications of Healthcare Data Breaches through Computational Technique. *Intelligent Automation & Soft Computing, 32*(3), 1763-1779.

Argun, U., & Daglar, M. (2016). View of Examination of Routine Activities Theory by the Property Crime. *International Journal of Human Sciences, 13*(1), 1188-1198.

Ayereby, M. P. (2018). Overcoming Data Breaches and Human Factors in Minimizing Threats to Cyber-Security Ecosystems. *Walden University Scholar Works*, 1-210.

Cheng, L., Liu, F., & Yao, D. (2017). Enterprise data breach: causes, challenges, prevention, and future directions. *WIREs Data Mining and Knowledge Discovery, 7*(5), e1211.

Cohen, L. E., & Felson, M. (1979). Social Change and Crime Rate Trends: A Routine Activity Approach. *American Sociological Review, 44*(4), 588-608.

Cremer, F., Sheehan, B., Fortmann, M., Kia, A. N., Mullins, M., Murphy, F., & Materne, S. (2022). Cyber-Risk and Cybersecurity: A Systematic Review of Data Availability. *Geneva Papers on Risk and Insurance, 47*, 698-736.

Dart, M., & Ahmed, M. (2023). CYBER-AIDD: A novel approach to implementing improved cyber security resilience for large Australian healthcare providers using a Unified Modelling Language ontology. *Sage Journals, 9*(1), 1-15.

Dean, N. S. (2023). Healthcare Data Breaches: Analysis and Prevntion. *California State University, San Bernardino*, 1-58.

Ignatowski, M. (2021). Contributing Factors to the Number of Individuals Impacted by Data Breaches in Healthcare Organizations. *Capitol Technology*.

Ismagilova, E., Hughes, L., Rana, N. P., & Dwivedi, Y. K. (2022). Security, Privacy, and Risks within Smart City: Literature Review and Development of a Smart City Interaction Framework. *Information System Frontiers*, *24*(1), 393-414.

Javoid, M., Haleem, A., Singh, R. P., & Suman, R. (2023). Towards insighting cybersecurity for healthcare domains: A comprehensive review of recent practices and trends. *Cyber Security and Applications*, *1*(6), 100016.

Karunarathne, S. M., Suxena, N., & Khan, M. K. (2021). Security and Privacy in IoT Smart Healthcare. *ORCA - Online Research*, *25*(4), 37-48.

Li, J., Xiao, W., & Zhang, C. (2023). Data security crisis in universities: identification of key factors affecting data breach incidents. *Humanities and Social Sciences Communication*, *10*(1), 270.

Mcleud, A., & Dolezel, D. (2018). Cyber-analytics: Modeling factors associated with healthcare data breaches☆. *Decision Support Systems*, 1-12.

Oloyede, A., Ajibade, I., Obunadike, C., Phillips, A., Shittu, O., Taiwo, E., & Kizor-Akaraiwe, S. (2023). A Review of Cybersecurity as an Effective Tool for Fighting Identity Theft Across the United States. *International Journal on Cybernetics & Informatics (IJCI)*, *12*(5), 31-42.

Oloyede, K. (2023). Impact Of Web (URL) Phishing and Its Detection.

Raghupathi, W., Raghupathi, V., & Sahara, A. (2023). Analyzing Health Data Breaches: A Visual Analytics Approach. *Journal of Applied Mathematics*, *3*(1), 175-199.

Reddy, J., Elsayed, N., Elsayed, Z., & Oser, M. (2021). Data Breaches in Healthcare Security Systems. *Journal of Information Systems*, 1-7.

Sardi, A., Rizzi, A., Sorano, E., & Guerrieri, A. (2020). Cyber Risk in Health Facilities: A Systematic Literature Review. *Journal of Sustainability*, *12*(17), 7002.

Seh, A. H., Zarour, M., Alenezi, M., Sarkar, A. K., Agrawal, A., Kumar, R., & Khan, R. A. (2020). Healthcare Data Breaches: Insights and Implications. *Journal of Healthcare*, *8*(2), 133.

Wilcox, P. (2015). Routine Activities, Criminal Opportunities, Crime and Crime Prevention. *International Encyclopedia of the Social & Behavioral Sciences*.