BOHR International Journal of Computer Science
2023, Vol. 2, No. 1, pp. 29–45
DOI: 10.54646/bijcs.2023.14
www.bohrpub.com

REVIEW

Data breach and privacy in the digital era

A. Oghene*

eProcess International, Ecobank Transnational, Accra, Ghana

***Correspondence:**
A. Oghene,
augustine.oghene@gmail.com

Received: 13 April 2023; **Accepted:** 27 April 2023; **Published:** 03 May 2023

The trend of data breaches globally is attracting concerns from business owners and the government of different countries. Notwithstanding the multiple pieces of legislation and strict requirements for reducing crime, there has been an exponential increase in data breach incidents. These crimes result in the loss of billions of dollars annually from small entities and large enterprises. Data breaches did not only start during the digital era with technological advancement; they started when individuals and organizations stored and maintained their data and records on-premises. However, as technology advances and computer systems become more accessible and affordable, coupled with poor management of sensitive documents, data breaches occur when individuals view other people's files without authorization. The rate of data breaches rose from 1980 to the early 2000s, giving rise to awareness of the canker. Laws and regulatory agencies such as Health Insurance Portability and Accountability Act (HIPAA) and Payment Card Industry (PCI) Data Security Standards were then established to guide sensitive data and the custodians. Regulatory frameworks developed as best practices to secure sensitive information are, however, not being implemented exhaustively, thereby not able to gatekeep satisfactorily. Data breach frequency is high in the digital era. The digital era exposes organizations and individuals continually to potential security breaches. Technologies like artificial Intelligence, machine learning, and data modeling make it possible to design algorithms and neural networks that help anticipate events and generate more data. Subsequently, fling open the gates to potential data breaches. In essence, it becomes absolutely necessary to consider the aftermath of information security and privacy in this digital era.

Keywords: data breach, privacy, information security, digital economy, cybercrime

Introduction

Data breaches and privacy issues have been a significant concern to individuals, organizations, and governments alike. The digital economy is impacted consequentially by data breaches. Research has shown that this crime will continue to grow. According to *Frontiers in Computer Science*, in May 2019, Canva experience a data breach that exposed the email addresses, usernames, names, salted/hashed passwords of 137 million users. The emerging digital technologies and control measures established by various organizations and countries to curb this crime is still not yielding the much-needed result. This paints a gloomy picture of ending data breach.

Data is classified as an essential part of an organization in the digital era. Data leakages threaten enterprises, including significant reputational damages and financial losses. With the increasing volume of data generated, the data breach rate is high compared to the past years. So, detecting and subverting the criminal act has become an extreme challenge and security concern for enterprises.

Cybercrime is primarily known to be responsible for data breaches once a system within a network is compromised; customers' data are taken from the environment. Technological advancement has unfortunately become a catalyst for criminals to invade people's privacy. These invasions happen consciously and unconsciously by individuals and organizations.

Interestingly, some organizations are making financial gains by trading with individual data. The organizations in this practice are referred to as shadowy companies. They collect an individual's medical history, for instance, from

social media accounts and build their behavioral profile and in turn sell the data to appropriate companies, this is done on the blind side of data owners. When these companies make a sale, more data is released, then it becomes continuous. Subsequently, the risk of future data breaches increases exponentially. The company data is sold to, cannot be held liable for data breach by the owners of the data, nor their management be tried for fraud.

Companies that trade data can be held liable for the breaches concerning data sold by them. Still, the least intrusive course of action is for the companies to be responsible for the breach. Meanwhile, the drawback is that by the time a breach happens, it is so far late. Over millions of records might have been exposed. Considering how data breaches and privacy occurs, preventive measures should be established and enforced to reduce the act to a negligible level.

Moreover, multiple data breach detection gap analyses revealed that uncovering data breach incidents could take months to identify. According to *Frontiers in Computer Science*, data breach reported by Verizon security research states that, hundreds of data breach incidents reported by multiple organizations show that more than a quarter of incidents had taken place unnoticed for many months, and one in ten are hidden for over a year. The Sony data breach incident occurred without visibility for almost a year, resulting in an estimated breach loss of one billion dollars. The aftermath of a data breach not only affect the competitive edge of an enterprise but an entire sector in the market space. It eliminates trust from the customer's perspective. When the incident is made public, it has a binding effect on the enterprise's bid-ask spread and trading volume performance—considering the data breach incident in Equifax. The announcement hurt its stock price. *Frontiers in Computer Science* state that enterprises should avoid reputation deterioration, and regulatory sanctions deliver non-quantifiable business benefits but significantly add value to business operations and success.

Research objective

1. Impact of data leakage threats to enterprise operations such as corporations and government agencies.
2. Increase in data breaches during the digital era, such as intellectual property, product specification, manufacturing techniques, personally identifiable information (PII), medical history of individuals, and credit/debit card information.
3. Organizational strategies must be robust to offer insight into how companies can move forward with data in an era of constant change. Regardless of the best information security policies and practices, an obvious threat to every corporate network is the internal staff's personal use of the corporate network.
4. Privacy implications of emerging technologies such as digital technology.
5. Optimizing data generated by legacy and emerging technologies while maintaining the boundaries of privacy regulations.
6. The organization, both private and public, does not adhere to data ethics. Focusing on security is inadequate as ethics is the new area businesses embrace with the digital transformation journey.

Literature review

In today's digital technology era, responses to data breaches are essential and complicated. The response becomes especially important when vital data like PII, intellectual property, credit/debit information, and individuals' medical information may be made accessible to an unauthorized person. Also, the response is important since it may negatively impact other specific needs of various stakeholders. Therefore, a suitable incident response plan should be executed with a multifaceted approach with unified coordination. As thousands of businesses change their operating model to improve efficiency, their operating environment becomes more volatile, and risk exposure may be heightened.

Although enterprises focus on protecting data that combines cybersecurity controls and incident response plans, hackers' threats are still growing exponentially. The organization must strengthen its incident response strategy to make it particularly challenging for adversaries to easily take billions of dollars worth of data from the organization.

Regardless of the fact that employees' security and data protection measures to secure the network from external threats are put in place, internal threats are often neglected. The organization considers all the firewalls, intrusion prevention, and detection systems, monitoring tools, and network sensor appliances with the best cybersecurity policies and practices. Meanwhile, a minor threat to the corporate environmental network may lead to a data breach through employees inadvertently executing malicious code through an embedded website.

Additionally, an enterprise can deploy the best security perimeter to safeguard its network. However, the compromise to the perimeter could be an insider. Therefore, cybersecurity policies and practices should encompass awareness training concerning phishing attempts, social engineering, malicious websites, and other cyber threats.

Notable data breaches and privacy issues led to the loss of hundreds of millions of dollars from many organizations with state-of-the-art technology, such as Yahoo, resulting in data breaches. Yahoo declared two data breach incidents in 2016. The first hacker compromised as many as 500 million user accounts toward the end of 2014. In

December 2016, Yahoo discovered another data breach resulting in over one billion user accounts. The data breach incident caused Yahoo to decline its acquisition by Verizon by less than the agreed amount. Verizon ended up paying 350 million dollars less than the initially planned sale amount. With technology connecting the entire world as a single entity, data is accessible from any location, and the perimeter is implemented elsewhere, securing vital assets from advanced hacking techniques. Disheartening data breaches result in direct and indirect cost factors that are profoundly responsible for every sector's sustenance and competitiveness—cutting across financial, reputational, operational, and the fulfillment of regulatory requirements. Also, as the impact of a data breach is exponentially increasing, government and regulatory agencies are introducing stiff new compliance and regulation laws such as General Data Protection Regulation (GDPR) and National Diploma in Business Studies (NDBS) to help make the existing controls more challenging to compromise.

According to the International Business Machines (IBM) research analysis on data breaches, the root cause of a data breach is linked to three categories: system glitches, including both IT and business process failures; human error, including negligent employees or contractors who unintentionally cause a data breach; and malicious attacks, which hackers or criminal insiders can cause. The research provides more insight into the percentage of incidents involving the three categories of data breaches. As many as 52% of incidents, such as a malicious attack, were recorded, compared to 25% caused by system glitches and 23% caused by human error. The trend in average total cost by the three categories of a data breach in 2020, measured in US dollars is displayed below (**Figure 1**).

Multiple factors, such as security technologies and practices, information technology environments, and involvement by third parties, influence a data breach's cost. According to the 2020 IBM report, the average total cost impact of the factors is estimated to be US$3.86 million. The complexity of security controls attributed to deploying technologies without expertise resulted in a data breach costing the enterprise an average of US$291,870. The report further revealed factors that mitigated the average cost estimate of a data breach across extensive testing of the incident response plan and business continuity management, reducing the average by an average of US$295,267 and US$2,78,697, respectively.

It further states that a cumulative breach in 2018, 2019, and 2020 of between 1 million and 10 million important records cost an average of US$50 million, more than 25 times the average cost of US$3.86 million for data breaches of less than 100,000 records. The 1–10 million record size of a breach experienced the most incredible growth rate, increasing by 22% from an average of US$39 million in 2018 to US$50 million in 2020. Then again, when the breaches involve more than 50 million records, the average cost is US$392 million, more than 100 times the average data breach cost. But the highest absolute cost increase was in breaches greater than 50 million records, leading to an average of US$350 million in 2018 and US$392 million in 2020 (1). **Figure 2** provides the average cost loss because of data breaches from 2018 to 2020.

Data breaches can be costly for an enterprise in multiple ways. Losing data, lack of customer trust, litigation, and difficult recovery from the incident of a data breach can also have dire financial implications. As the incident's financial implications keep increasing, there is no end in sight. The health sector, for instance, experience data breaches more than other sectors due to the volume of records. It is a highly regulated industry with many regulatory burdens, but most importantly, it requires conforming to regulatory requirements. Additional costs come with ransomware, most notably deciding whether to pay to retrieve data. The average cost to rectify the impacts of the most recent ransomware attack is US$ 732,520 for organizations that do not pay off the ransom, rising to US$ 1,448,458 for entities that do.

However, ransomware data breaches happen mainly in the private sector compared to the public sector. As much as 45% of the public sector was hit by malware such as ransomware last year, compared to a global average of 51% and a rate above 50% in the media, leisure, and entertainment industries. However, ransom data breach incidents were covered by insurance, thereby taking up the cost (2).

Besides an enterprise's financial losses, other potential penalties include unsuitable cybersecurity controls implemented, legal fees, discounts, incentives to retain customers, and identity theft protection control investments to secure customers and the enterprise's employees. Research has shown that data breaches financial impact has devastated many enterprises of all sizes, including tiny- and medium-sized businesses. Furthermore, the study established that organizations with fewer than 1,000 employees were mainly hit, resulting in losses estimated to be millions of dollars on average. The financial losses and other pains arising from data breach crime cannot easily be isolated when an enterprise is attacked after many months and years. The threat remains for many companies, and the financial costs are sizable (3).

Data governance

Globally, as enterprises in different sectors such as financial institutions, media, health care, telecommunications, and manufacturing companies generate massive quantities of data, the fear of data breaches and privacy has drawn the attention of the government, regulatory agencies, and private agencies to act. Although each enterprise has a governance strategy, the risk pattern and regulatory landscape demand an overhaul of the governance methodology. Data governance is the data management function designed to guarantee the quality, integrity, security, and usability of the data collected

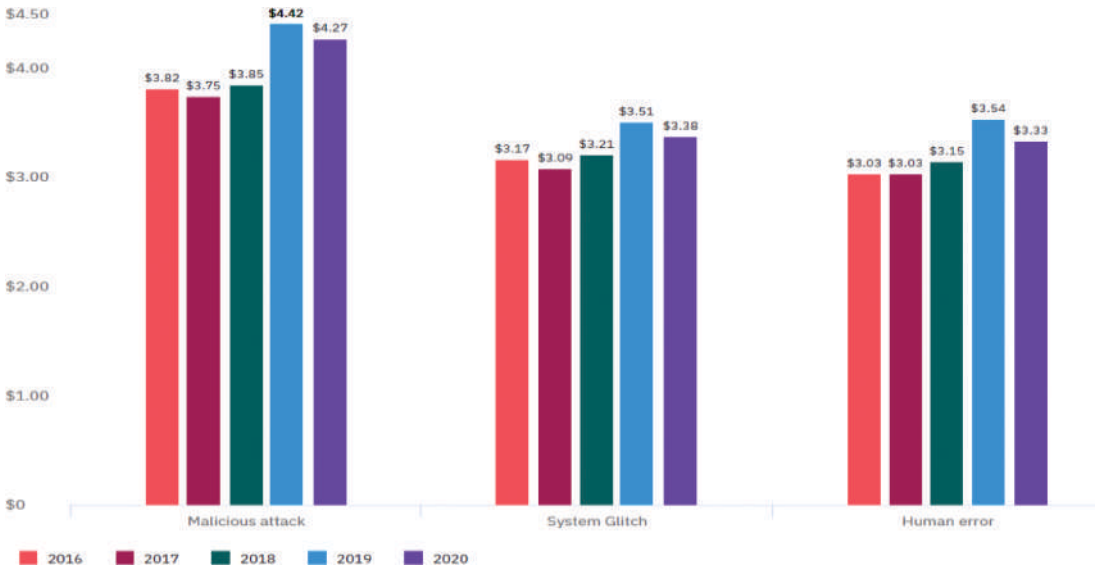

FIGURE 1 | The trend in average total cost for the three categories of a data breach in the year 2020 is measured at US$.

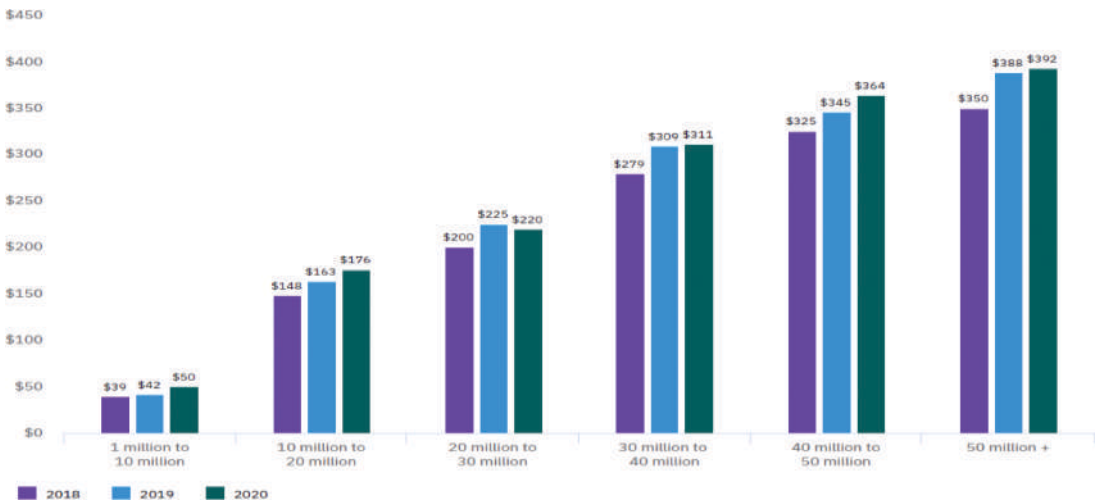

FIGURE 2 | Provide the average cost loss because of data breaches incidents from 2018 to 2020.

by an organization (4). Lack of governance increases the risk of enterprise and the absence of quality data, which resulted in the financial crisis of 2008 (5). Data governance is a set of processes that ensure that essential assets are formally managed throughout the enterprise. Operationalizing data governance provides the assurance that data can be trusted and someone can be held accountable for any adverse event that occurs, impacting the quality of data (1). Also, data governance is achieved through a cultural change that spans the whole organization across functions, units, domains, and roles (6).

More so, the data life cycle ensures data governance and emphases, making the data available to all stakeholders in a state that is readily accessible and in conformity with regulatory standards. Regulatory standards often intersect with industries like health care, government, and company rules and behavior codes. It becomes imperative to the enterprise; stakeholders get a single view of their data within the organization (4).

The benefit of data governance is tailored to enhance trust in the data for all stakeholders and businesses. This is because reliable data is a vital precondition for management to make the right decisions. It applies to all enterprises, regardless of their size. Data governance cannot be ignored and should be demonstrated to establish its benefits to the enterprise. According to Ref. (4), one of the recurring questions that our enterprise customers have is about what best practices and policies they should put in place to manage the classification, discovery, availability, accessibility, integrity, and security of their data. However, Ref. (5) states that the ideal policy as a data-driven governance approach is one where governance is established at the data level, unlike the system level through different platforms wholly apart from the data. In

this case, implementing governance at the data level allows the database logic to manage constraints at the database level.

In contrast, establishing governance at the code level requires developing specific code. It became unavoidable that enterprises must make vital decisions concerning their data. This decision should be based on how the data is saved, archived, backed up, and protected and followed by auditing their security policies and reviewing the security in alignment with new security trends to quickly develop a strategy to subvert threats resulting in data breaches and privacy issues.

Moreover, Ref. (5) states that data governance should conform to basic principles such as confidentiality, integrity, and availability. These keep the bad guys from having easy access to data. The governance approach will increase the trustworthiness of enterprise data for the benefit of the business. Achieving this requires accountability by defining an operating model that allocates ownership and accountability across all data boundaries. The need for an access control mechanism is significant in the digital era, where data breaches and privacy issues are more prevalent. The mechanism will facilitate the governance of data to restore assurance and trustworthiness. Classification and access control should be applied to data as part of the governance process to derive business benefits (4). Data governance has gained prominence since the inception of GDPR and California Consumer Privacy Act (CCPA) type regulations (5). A sound data governance model should cover data security to avoid putting data at risk. This means the data must be secure and trustworthy. Among the various models, the Mark-Logic security model is the best. It provides fine-grained authorization with role-based access control, allowing users to be assigned multiple roles; the roles are created with permissions and privileges. The users can only perform certain activities with document-based permission, such as read, insert, and update, and the privileges allow the user to control what actions they can perform (7).

Throughout data governance, techniques for improving compliance and mechanisms for doing so efficiently and effectively are deliberated. The outcome of such deliberation provides more insight into the governance and security of data. Also, Ref. (4) states that the increase in data and its availability have resulted in the desire and need for regulations on data collection, access, and use. Regulations such as Health Insurance Portability and Accountability Act (HIPAA) have existed for quite some time, while other regulations like GDPR in the European Union and CCPA are in force in the United States. Both regulations apply to multiple companies.

Data strategy

Unlike before the evolution of information technology, many enterprises generate more data in the digital era, leading to digital technology. Developing a data strategy requires the enterprise's top hierarchy's commitment to understanding that data is a corporate asset that must be properly managed and protected like another critical asset. According to Ref. (8), harnessing data as a corporate asset requires a mix of soft skills required to build sustainable strategies and manage change and hard stuff, which applies a portfolio of data management tools and techniques to ensure the delivery of consistent, high-quality data that is aligned with business strategies and initiatives. Also, Ref. (9) states that, with a robust data strategy, the enterprise can respond to new opportunities faster and improve the customer's experience. The enterprise should embrace a collaborative and data-driven culture that makes it easy for its leadership to make well-informed decisions.

Moreover, every enterprise must imbibe a vital point in data management: identifying the difference between enterprise and local data requirements. And Ref. (9) mentioned that most organizations whipsaw between these two extremes. Regardless, astute managers foster a dynamic interplay between the two polarities, embracing both simultaneously without getting stuck at either end. Many have a centralized approach to data management while at the same time sustaining data integrity and quality.

Also, there are multiple enterprises with fragmented data. These organizations must implement processes and controls to help standardize their data management. Empowering a data-driven culture within the enterprise ultimately makes everyone understand the need to protect and extract value from it. According to Ref. (10), recognizing success depends on empowering internal users beyond the data engineering and analytics teams to buy into the data-driven culture.

A fundamental element of data strategy is designing and executing a data governance program that benefits the enterprise model of generating and managing data. However, not every organization that grows data at a breakneck pace has adopted the data governance model to manage their data. But increased regulatory pressure, such as GDPR and the CCPA, has created a sense of urgency among organizations. They are prompting executives to explore modern data governance programs.

Furthermore, the governance of data is beneficial to enterprise businesses and operations. It makes identifying personal data and validating data within enterprise ecosystems for compliance easier. In some cases, organizations find it challenging to develop a comprehensive data catalog of all data within the entire organization and its usage across business functions, geographies, and ecosystems. All these are addressed by regulatory compliance. In a nutshell, Ref. (7) states data governance is an overarching strategy to validate that the data is clean, accurate, usable, and secure, resulting in data integrity and good governance, which are paramount to the enterprise's ability to make the right decision for the entity's benefit.

There must be alignment across the following: the people, processes, and technology to facilitate the identification,

classification, and documentation of information about the organization's assets. These can be achieved when the various employees from different business functions with clearly defined roles and responsibilities understand the data generated within the ecosystem and are empowered to provide insightful information on the data for business decisions.

The correct data governance processes are also essential for an enterprise management team to conform to the requirements of GDPR, CCPA, and other regulatory agencies. The requirements are made strict as data breaches and privacy issues increase exponentially each year with financial loss. Organizations do this review by making amendments to the contract appropriately, extending the privacy and security standards to third parties with access to their data and the data they have monetized. However, Ref. (7) mentioned that enterprises should respond to breaches and privacy issues within their organizations and with a third party if need be.

Also, the strategy should be designed to combat attacks and protect data. This is essential as multiple organizations have relocated to network enterprises to enhance efficiency; there is a threat to the network environment that is now more volatile, increasing the risk of exposure and a heightened risk of compromise. With this, enterprise strategy must focus on protecting data to combine network security measures and incident response programs. Building a data strategy should be considered from a technological perspective before transitioning it into reality. Beyond that is an understanding of the business-side challenges with the strategy. But new business models can be created through proper data management and an excellent data strategy. This can only be possible when all the employees are aligned and there are clearly defined processes for data management.

Moreover, the focus should be on data management instead of strategic reasons. Data management cuts across the following: development, execution, and supervision of plans, policies, programs, and practices that conform to defining controls and data protection and increase data value across their life cycles (8). Performing data management makes the database continue running with optimization and daily data backup. Unlike a data strategy developed to drive a data management program strategically, this requires a plan for maintaining and enhancing data quality, integrity, data access, and security while mitigating identified risks.

Research methodology

The quantitative descriptive research study provides systematic information about the research. It revolves around describing, explaining, and validating the findings. The method is designed to depict participants in a very systematic and accurate manner. Analysis of the collected data will be approached in a structured manner. A closed-ended questionnaire will be developed as the tool adopted for data collection in this research. The questionnaires will be distributed online to individuals in different sectors. In addition, with the internet, the respondents can receive the questionnaires on their intelligent mobile phone or portable device through emails.

The data collected from all the respondents will be subject to in-depth analysis to eliminate anomalies. The research approach is based on inductively exploring research involving data breaches and privacy across different sectors.

The study uses both qualitative and descriptive types of methods. As a result, the analysis used various keywords relating to data breach trends to ascertain the impact on the organization. The approach is tailored to address all the research questions. They considered that data breaches in the digital era had taken on a different dimension, unlike in the past. The dynamic of a data breach leverages multiple data sources created through technological advancement. Individuals and organizations have not realized the need to secure data from both internal and external threats.

Data collection

Data collection is primarily from individuals working in sectors such as finance, telecommunications, health, and government agencies. Pulling data from multiple sources provides research and analysis with accurate information to address research questions. Most importantly, data sources from across all sectors help quantify the consistency of the findings concerning data breaches in each sector compared to others. All the data will be collected from participants who received the questionnaire link and shared their feedback after answering the questions. The questionnaire and feedback are based on the data breach factor. This harmonizes with the literature review, considering all the types of data breaches and privacy challenges confronting the sectors.

Analysis of the finding

Respondents to the questionnaire share their feedback based on their perception of the data breach and privacy incident in their organization and the industry. Although there are various ways of conducting the research, this was executed by preparing a questionnaire and making it accessible online through a link. The respondents' feedback is represented in graphs such as pie charts and bar charts. These helped put their feedback into perspective. The analysis covers substantially more industry breakouts relating to data breaches initiated through attacks and actors. Also, the feedback from each respondent makes the analysis arguably the most comprehensive analysis of data breaches (**Figure 3**).

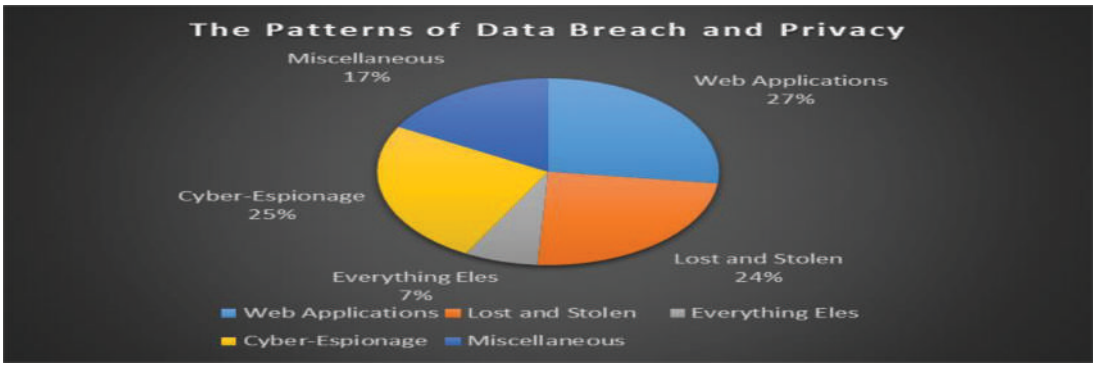

FIGURE 3 | The patterns of data breach and privacy.

Data breach and privacy manifest in multiple patterns through web applications, miscellaneous, cyber-espionage, lost and stolen items, and everything else. The respondents' feedback indicates web applications are mostly used to commit data breaches and privacy incidents. However, 27% of web application attacks arose from industry data breaches and privacy issues. It occurred through attacks against the code of the actual web application, such as exploiting code-based vulnerabilities.

Further, 25% of cyber-espionage attacks are responsible for data breaches. This cyberattack looted classified, critical data to gain an advantage over a competitive company and government entity.

Also, the incident of data breaches occurred through the following: 17% of miscellaneous attacks lead to it, and 24% of lost and stolen date are responsible for data breaches. These incidents include when an asset or piece of data might have mysteriously disappeared. It could also be accidental, and 7% of everything else results in data breaches. These incidents are phishing or financially motivated social engineering, where attackers try to commit fraud through email (**Figure 3**).

There are actors behind cyber-incidents that result in data breaches. Data breaches take time to pull off. Successful data breaches have a substantial impact not only on organizations but also on millions of people. According to the respondent's feedback, 24% of organized criminal groups and internal actors are firmly behind data breach incidents. While 21% are external actors exploring ways to infiltrate and cause harm to the organization by carefully looking for weaknesses, 13 and 18% indicated that multiple parties and partner actors play a significant role in data breaches manifesting (**Figure 4**).

The execution of data breaches is based on the vulnerable state of the entity's infrastructure ecosystems and the employee's awareness concerning protecting themselves from cyberattacks. The respondents' feedback shows that 26% of malware attacks result in data breaches. The malware attack is a popular cyberattack in which malware executes unauthorized events on the personal target device. The attack is possible when a cyber-actor develops malicious software installed on someone's system without the person's knowledge to gain access to personal information for financial gain.

Social engineering and hacking attacks are 25 and 23% of the most exploited tactics for causing data breaches, respectively. Social engineering attacks take time to monitor their targeted victims and gather vital information before launching. Hacking attacks look for ways to access their victim's computers to steal information, install malware, launch a ransomware attack, and sometimes enter the network through unsecured systems. Each attack led to data breaches and privacy issues.

Misused unauthorized users and errors are 17 and 9% of the tactics exploited by cybercriminals to launch attacks causing data breaches, respectively (**Figure 5**).

However, 25% of respondents' feedback shows "financially motivated" as one of the features. The majority of cyberactors are financially motivated by criminal acts. And the financial sector remains the prime target for a financially motivated hacker.

Further, 23% of respondents' feedback shows phishing and stolen credentials are both standard features of data breaches. Phishing attacks were associated with a high rate of stolen credentials and social attacks. Many credential thefts stemmed from phishing attacks. However, 18% of respondents indicate that a proportion of malware incidents are Ransomware. Ransomware is a common malware breach.

Moreover, 11% of respondents mentioned web applications as a standard feature of data breaches. Cyber-actors prefer exploiting web applications where their software is poorly written because the software leaves behind holes, which can be an entry point to cause harm to data (**Figure 6**).

Every individual is a target of cyberattacks, regardless of their personality. This is because cyberactors launch attacks randomly for financial gain. In fact, 33% of the respondents indicated that data breaches involved significant business victims. This attack occurs through various cyber vectors, such as social engineering, malware attacks, hacking, and stolen or lost credentials.

Further, 56% of respondents state that individuals with personal data are the prime target of cyberattacks. The attacks

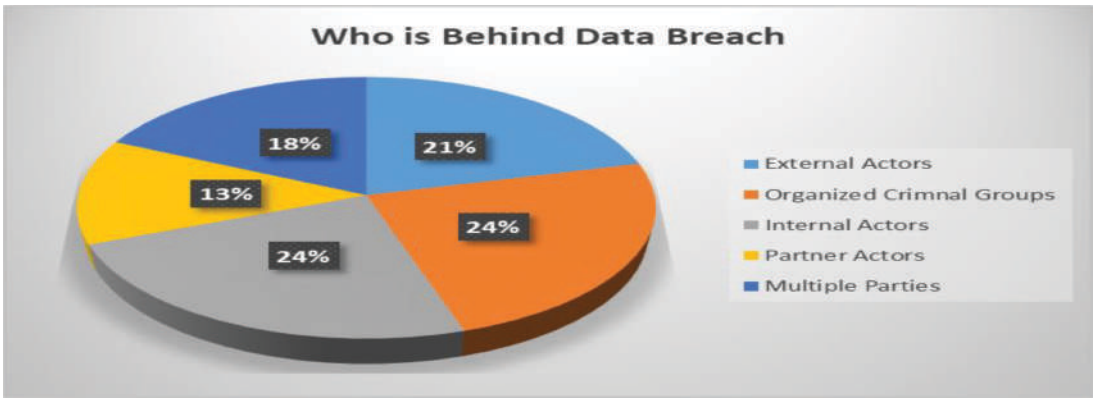

FIGURE 4 | Who is behind the data breach?

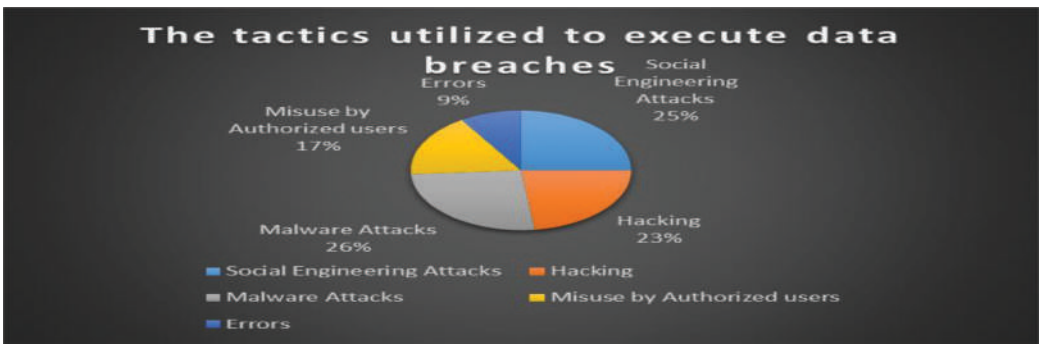

FIGURE 5 | The tactics utilized to execute data breaches.

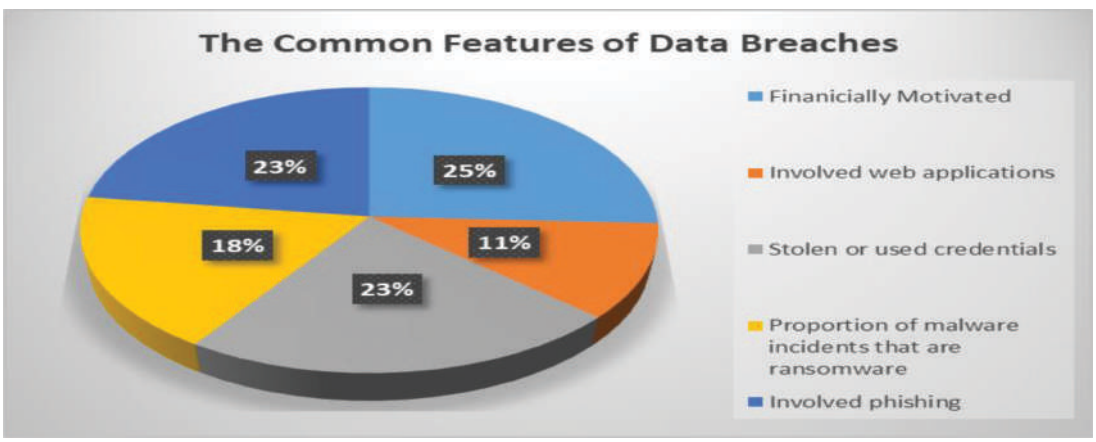

FIGURE 6 | The common features of data breaches.

leading to data breaches with individuals are primarily through social engineering tactics—11% of respondents' breaches occurred in days or less (**Figure 7**).

The success of cyberattacks on any organization depends on how secure the operating environment is against cybercriminals. Many organizations' vulnerable state allows exploitation tactics by cyberactors to break into the network to steal data and sometimes manipulate the data for easy financial gains. Fifteen respondents indicate that organizations lack time to keep up with vulnerabilities and potential attacks. Several small, medium, and large organizations do not have processes to manage their infrastructures' vulnerabilities, making it easy for hackers to find their way into the ecosystems without notification. Sometimes, there are no technological tools to secure the environment. Hence, eight respondents' feedback mentioned the absence of technological tools to identify vulnerabilities and correlate potential security issues across the enterprise—these are common with enterprises in different sectors.

Financial investment in security is not receiving adequate attention, allowing the environment's vulnerable state to persist. Eight respondents indicate a lack of inventory is

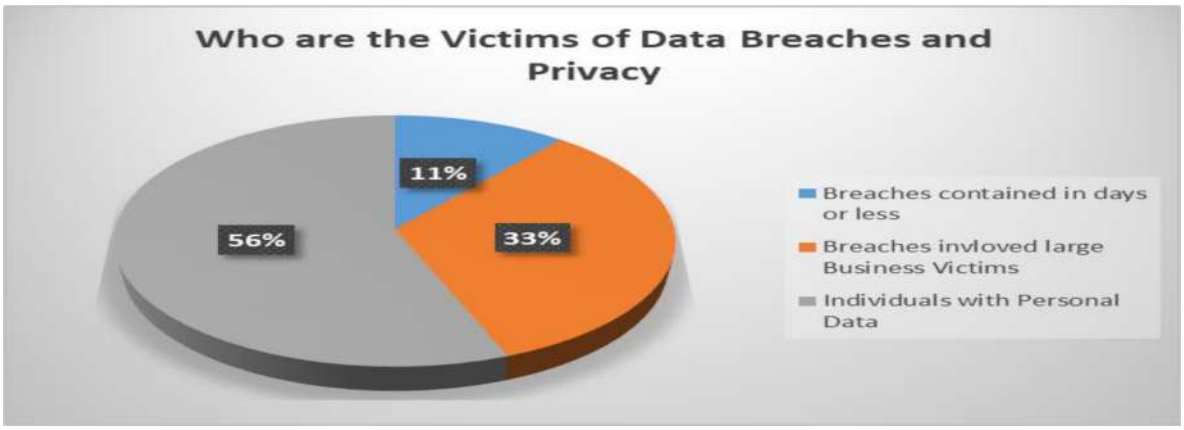

FIGURE 7 | Who are the victims of data breaches and privacy?

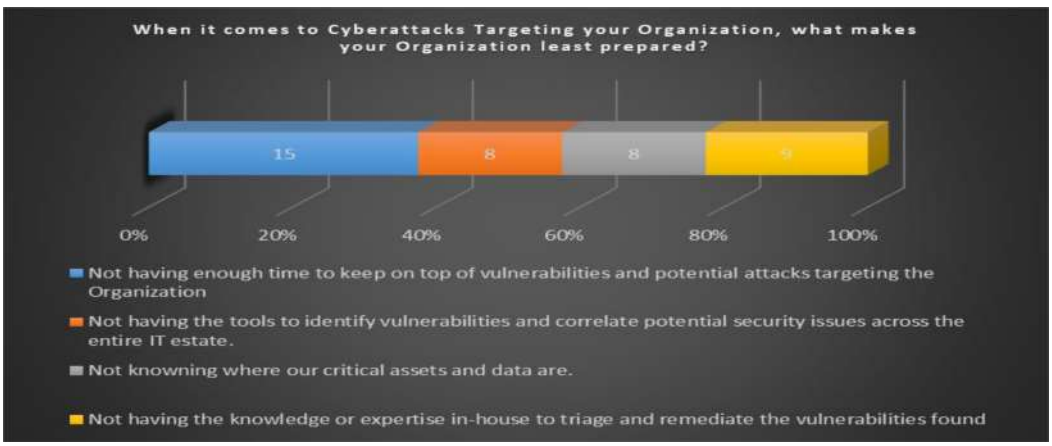

FIGURE 8 | When it comes to cyberattacks, target your organization; what makes your organization the least prepared?

FIGURE 9 | Threats to small organizations.

why cyberattacks are successful. Regardless of their size, organizations do not have a complete inventory of their assets. With this, only limited assets are within the scope of cybersecurity controls. Ninety respondents mentioned that experts' skills are insufficient, making it challenging to identify vulnerabilities and their remediation. Most importantly, implementing cybersecurity controls within the enterprise and beyond is a significant setback for many organizations (**Figure 8**).

Cyberattacks are everywhere, regardless of the size of the entity. Cybercriminals adopt multiple threat vectors to break into organizations. Most small entities have not protected their environment against threat vectors. Only a select few have invested in cybersecurity controls to secure their assets.

However, 17% of respondents stated that stolen credentials are used to steal individuals' PII and occasionally as an entry point into the organization through hacking and other vectors. Further, 22% of respondents show that the

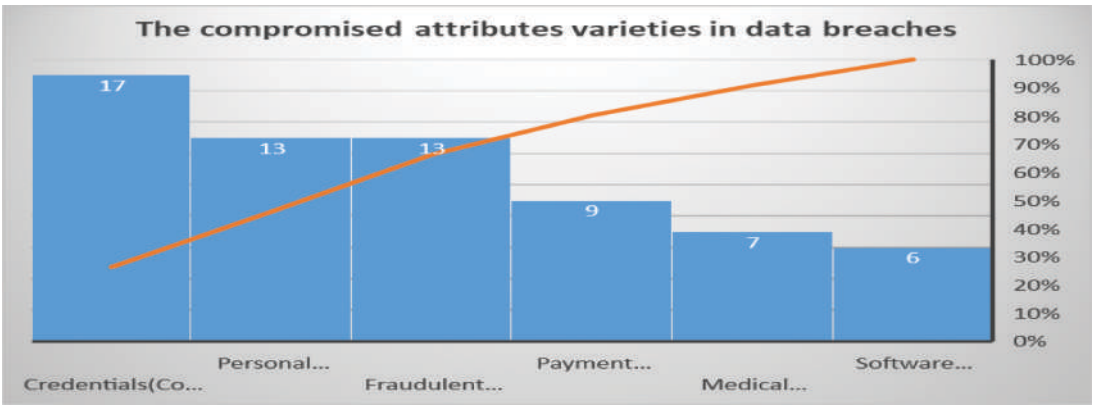

FIGURE 10 | The common attributes varied in data breaches.

phishing (social) vector is a significant threat to small entities. A phishing attack is a social engineering attack mainly used to steal sensitive information or data, such as login credentials, credit card information, and other vital details, by impersonating oneself as a trustworthy entity and communicating through email or instant text messages.

Moreover, 17% of respondents mentioned that privilege abuse is a means of launching cyberattacks on small organizations. When employees use their login privileges for new devices, access to the old ones is retained as their significant enterprise changes. This is a common practice in both small and large entities, unfortunately. Regularly correlating current privileges and roles in the enterprise with the actual business needs of rightful employees is imperative. Failure to actualize this practice may result in a data breach and privacy issue.

However, 16 and 15% of respondents indicate both export data (malware) and spyware (malware) threaten smaller organizations. Also, 13% of respondents mentioned brute force (hacking), another threat vector, as a threat to the small entity. Export data and spyware are malware types of vector attacks designed to cause harm to personal devices; they are contractions for malicious software. Simultaneously, brute-force attacks are launched through a trial-and-error approach to guess login information and encryption keys and locate a hidden web page. These threat vectors eventually result in data breaches (**Figure 9**).

Cyberthreats launched by the actors through various vector tactics bring about inevitable compromise known as personal (confidentiality), Credentials' confidentiality, payment confidentiality, fraudulent transaction integrity, medical confidentiality, and software installation integrity. These attributes indicate a compromise of the two principles of cybersecurity known as confidentiality and integrity. However, 95% of the respondents state that credentials' confidentiality is the most significant compromised attribute in data breaches. Further, 75% of the feedback mentioned that personal and fraudulent confidentiality are attributed to data breaches in the cyber-attack trend. In fact, 50% of respondents classified payment attributes as one of the types of data breaches after a threat attack. Moreover, 45 and 35% of respondents mentioned payment confidentiality and software installation integrity as part of the attribute varieties common in data breaches, respectively. A booming threat attack trend brings about the compromise of specific attributes at the end of the attack (**Figure 10**).

The threat vector phishing attack is a social engineering tactic to launch fraudulent attacks on their target victims to steal sensitive information belonging to individuals and organizational entities. Attributes of phishing attacks are emails initiated by the actors. They make unrealistic threats or demands in the form of scams, wrong spelling and grammar, malicious URLs, requests for sensitive information, and malicious links provided randomly to all users. Threat actors exploit all these red flags to bring about data breaches. Sixteen of the respondents state that personal and credential data are among the types of data compromised in a phishing breach. Twelve respondents' feedback shows bank data and classified data are varieties of compromised data in a phishing breach. Nine and four respondents' feedback stated medical and system data were compromised in phishing data breaches. Data is compromised when a threat vector such as phishing is launched, depending on the entity (**Figure 11**).

A malware vector attack is a cyberattack trend where malware initiates unapproved actions on its target victim's device through malicious software. The malicious software includes several types of attacks: ransomware, spyware, and command and control. The attributes after a malware attack include slowness, making it difficult for the operating systems and programs to start up; when the systems suddenly display a blue screen known as the blue screen of death, they frequently display a lack of storage space; and unwanted program pop-ups come up frequently. As many as 44% of the respondents indicated that exploiting the vulnerability is one of the malware varieties of data breaches. However, 28% of respondents said ransomware is the top malware variety taking place in every data breach left behind after a successful threat vector attack. The Trojan is another type of malware that 22% of the respondents identified as a top malware variety manifested in data breaches as evidence of the attacks.

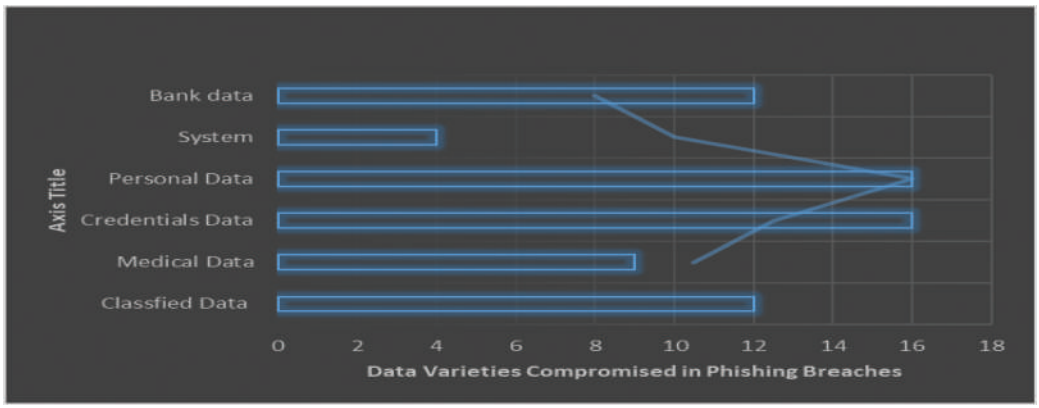

FIGURE 11 | Data varieties are compromised in phishing breaches.

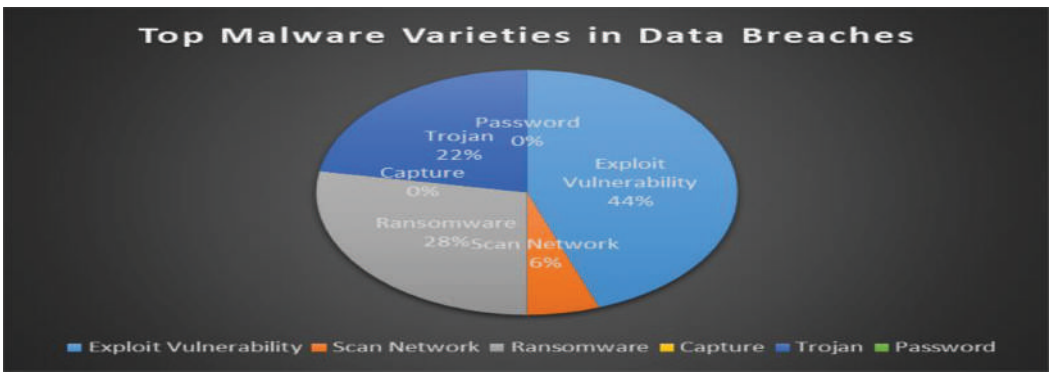

FIGURE 12 | Top malware varieties in data breaches.

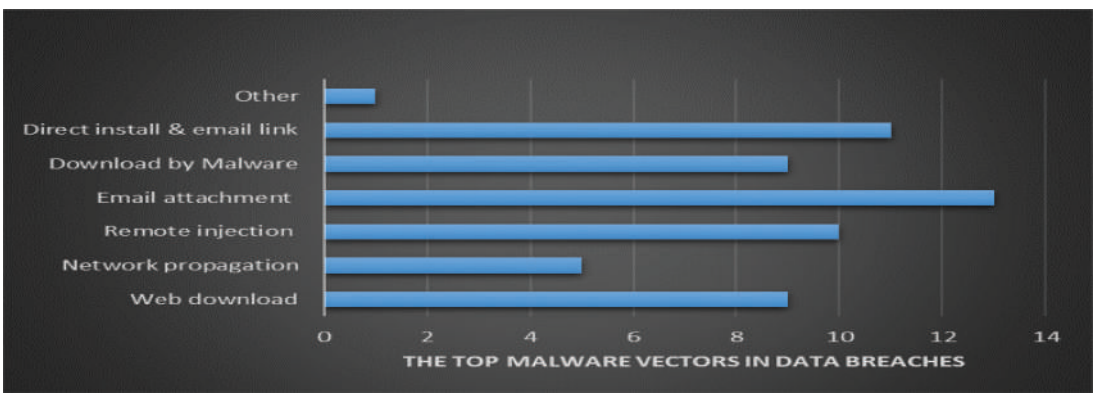

FIGURE 13 | The top malware vectors in data breaches.

As many as 6% of the respondents classified networks as malware varieties of a data breach (**Figure 12**).

Malware vectors pose threats to individuals and business entities. They are malicious codes such as viruses or worms that primarily launch themselves into a system or platform. Malware is distributed through social engineering (Trojans), Browser exploits, document/application exploits, and remote code execution in operating systems. When malware infects a system, it spreads by replicating itself, like a biological virus infecting host cells. In this scenario, when an infected device is connected to a network through open Wi-Fi and a memory stick, the virus is propagated across the network, and the memory becomes infected. Thirteen respondents state that email attachments are the top malware vectors in data breaches. This is launched via phishing malware. Ten feedback states that remote injection malware is among the top malware vectors in data breaches. This occurs when an attacker exploits an input validation flaw in software to launch and execute malicious code.

In most cases, the code is injected into the language known to the targeted application and initiated by a server interpreter. Nine respondents express that both downloads

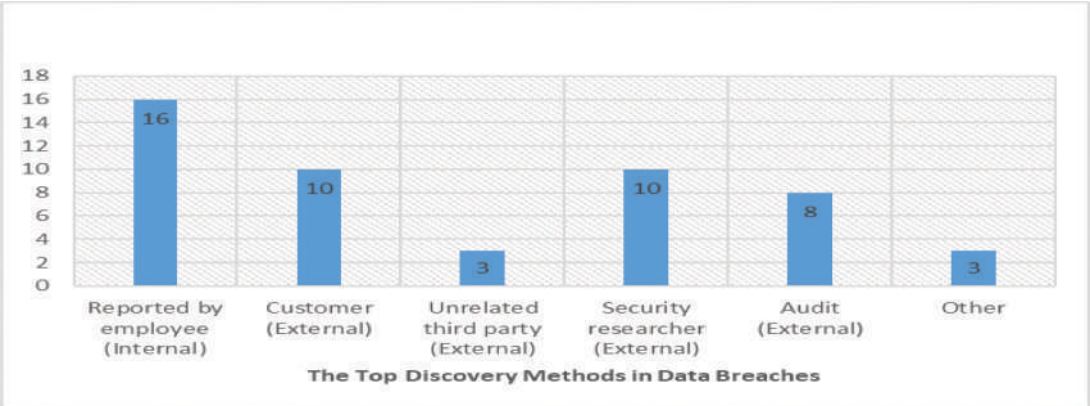

FIGURE 14 | The top discovery methods for data breaches.

FIGURE 15 | Top threat action varieties in cyber incidents.

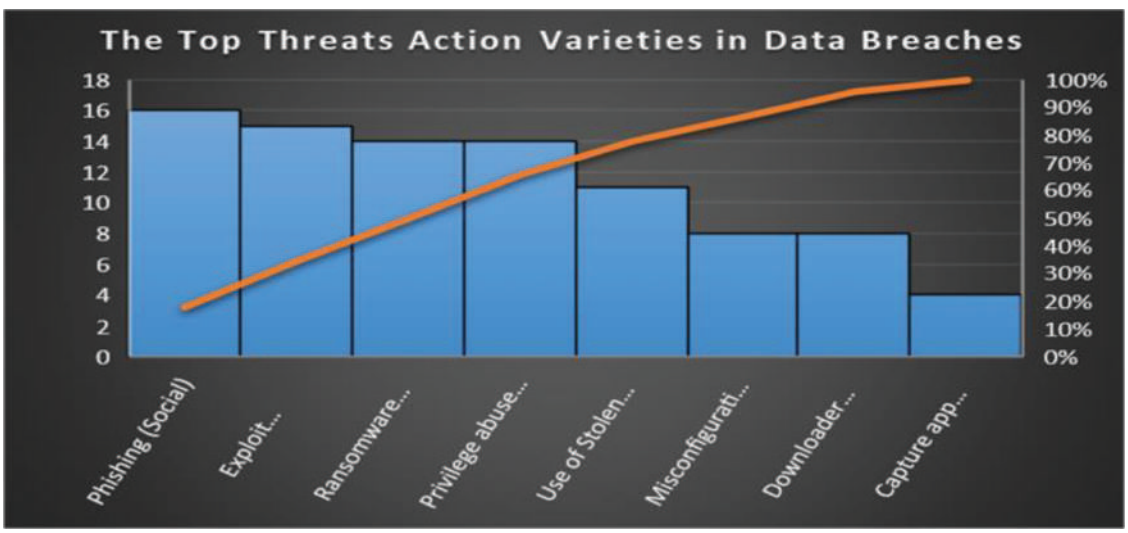

FIGURE 16 | The top threat action varieties in data breaches.

by malware and web downloads are top malware vectors resulting in data breaches. These malware vectors create data breaches when malicious links are clicked intentionally or unintentionally, resulting in the virus's download. Similarly, the web download act brings about viruses that cause harm to data. Five respondents mentioned network propagation as a top malware vector in data breach incidents (**Figure 13**).

Data breaches and privacy incidents are discovered in multiple ways and reported by different stakeholders. The incident was identified through technology controls

implemented to safeguard organizations against threat actors. Sixteen respondents state that internal employees report data breaches, which constitutes the majority method of reporting the incident. Ten comments received mentioned that external customers and external security researchers can raise a red flag when a threat attack occurs, leading to data breaches. Eight respondents expressed that external auditing is the top discovery method for data breaches. Three pieces of the feedback mentioned external, unrelated third parties, and others constitute part of the discovery method of confirming data breaches in an entity (**Figure 14**).

Threat action initiates attacks in multiple ways to cause damage to individuals and organizations, resulting in data breaches and privacy issues. The success of the attacks depends on the degree of cybersecurity controls implemented in the enterprise ecosystems. This threat actor will exploit cases where the vulnerabilities are not remediated and remain through cyber vectors. Therefore, cyber-incidents are triggered by the following threat actors such as Dos (hacking), DoS (malware), exploit the vulnerability, phishing (social), ransomware (malware), and use of stolen credentials (hacking). However, 23% of respondents' feedback received stated DoS (hacking) constitute threats action varieties of a cyber-incident. Further, 21% of feedback states that ransomware (malware) forms part of the top threats leading to cyber incidents. Moreover, 16 and 15% of respondents' feedback classified phishing (social) and exploited vulnerability threat action launched that led to a cyber-incident. At the same time, 14 and 11% of respondents state DoS (malware) and use of stolen credentials threat actions are responsible for cyber-use—also part of the top threat action varieties of cyber incidents (**Figure 15**).

Threat action varieties allow making a deep dive into bad guys' tactics. This figure provides a concept of what action varieties increase data breach incidents. There are various threatening plays that play a large part. And this ranges from phishing, exploitation of the vulnerability, ransomware, privilege abuse, use of stolen credentials, misconfiguration, downloading, capture app data. The respondents' feedback shows that phishing, exploitation, ransomware, and privilege abuse are the top threats in data breaches. The feedback represents different percentages depicted in the bar-chart diagram, with Phishing (social) indicated as 90%. Other threats like stolen credentials, misconfiguration, downloads, and downloaded app data fall at the chart's lower level. However, their threats result in data breaches (**Figure 16**).

Cyberattacks occur regardless of the size of an organization, and if the organization does not have sophisticated technology such as tools implemented to send out a notification once the ecosystem is compromised, the actors will remain in the network scanning through the entire operating environment looking for financial data to manipulate for their benefit. Though the actors have sophisticated tools to break into an enterprise network, it can be straightforward where the cybersecurity controls implemented are not adequate to safeguard the entity against the external intruder. In some cases, when a threat action compromises the enterprise network, it takes several days and time for some organizations to know they have been compromised. While others depend on the adequacy of their cybersecurity control implemented to get notification of cyberattacks.

Nine respondents state it will take hours for organizations to discover they have been compromised, leading to data breaches. The eight comments received mentioned it would take minutes for the organization to be aware of the data breach. Six respondents' feedback stated it would take weeks. Others express their feedback as days and months for the organization to get notified of the incident. Most scenarios of compromise resulting from a data breach depend on the mature level of cybersecurity management process and controls as well as the information security program directing the affairs and prioritization of protecting the environment (**Figure 17**).

Regardless of size, every organization risks experiencing a data breach. Cybercriminals continuously exploit the path of least resistance. As such, back doors opened by small entities are the target for exploitation. Since the data landscape is easy to navigate, they are impacted differently. Researchers say more than half of all small businesses incurred a data breach in the past year. This crime resulted in a large percentage of them going out of business quickly. Cybersecurity management is a bottom-line issue for small business owners, and overcoming the associated challenges is a significant difference between a successful business and financial loss. However, 90% of respondents state it will take months for SMBs to discover a data breach has taken place. Depending on the business context, it may be less than a month; hence, 60 and 50% of respondents mentioned weeks, days, and hours for SMB to discover a data breach and privacy event. However, 10% of respondents agreed it would take a minute and a second for SMB to discover a data breach event. The maturity of SMBs concerning the degree of sensitivity to cybersecurity activity in the organization will expedite the discovery of a data breach when it occurs (**Figure 18**).

Cyber-actors look for the least path to exploit to actualize their goals. Entities proactively monitor their operating environment for suspicious activities. This suspicious activity could be taking place in their network. The ability to quickly recognize these activities is paramount to narrowing the data breach's source and nature, making it easy for the security team to intercept the threat and reduce the damage.

How holistic the security controls are implemented determines the degree of safeguarding for each asset within the organization. However, 15% of respondents mentioned databases and web applications are easily compromised assets for exploitation. The vulnerable state of the assets is the determining factor that leads to compromising the assets. In cases where a zero day attacks an asset with a

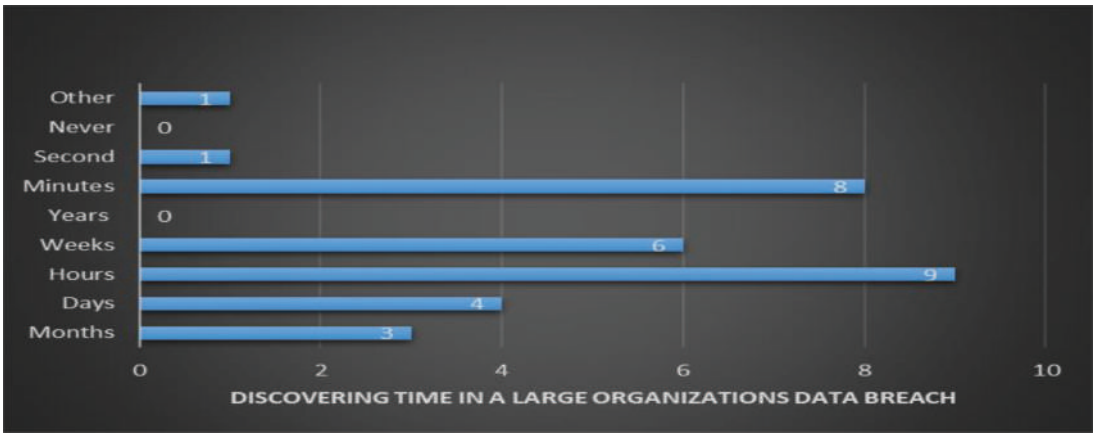

FIGURE 17 | Discovering time in a large organization's data breach.

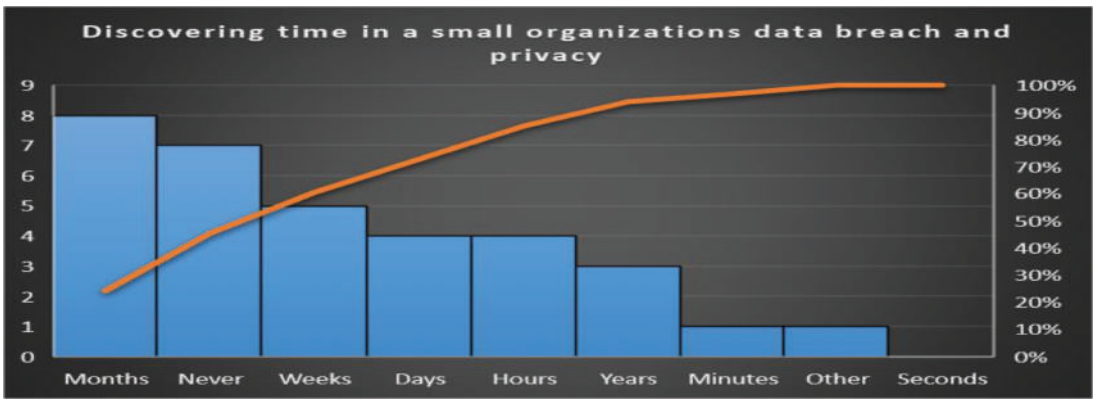

FIGURE 18 | Discovering time in a small organization's data breach and privacy.

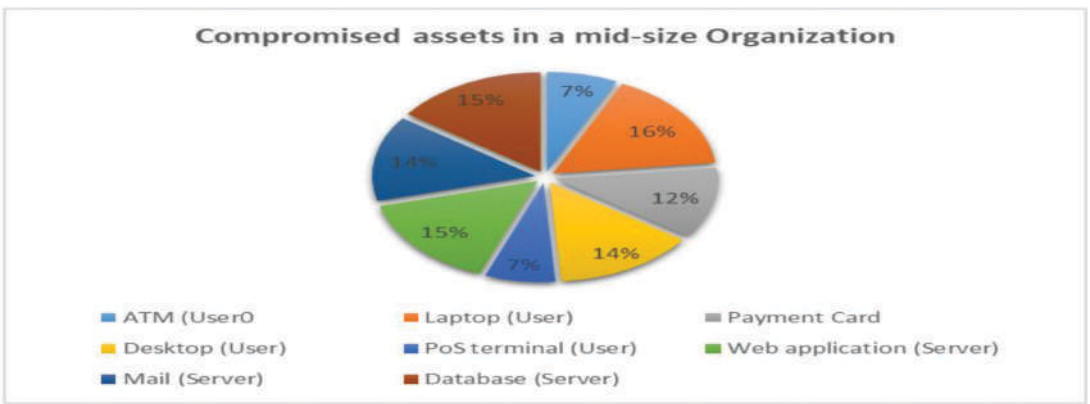

FIGURE 19 | Compromised assets in a mid-size organization.

new vulnerability, it will be compromised. Moreover, 16% of respondents mentioned laptops are easily compromised when the device is stolen and not properly secure. The information it contains is accessed, becoming a means of breaking into the enterprise. Further, 14% of respondents indicated desktop and point of sale (PoS) terminals are compromised organizational assets. Simultaneously, 12% of respondents agree that a payment card is a stolen asset to compromise a mid-size organization (**Figure 19**).

When assets within an enterprise are compromised, it leads to a data breach incident. This can happen when the entity's implementation of cybersecurity controls does not cut across the entire ecosystem—it allows some of the assets to be in a vulnerable state for exploitation by cybercriminals. In a large enterprise with a footprint spread locally and geographically, thousands of endpoint devices connected to the infrastructure ecosystems with few linking third-party organizations are an exit-point for different reasons. All the

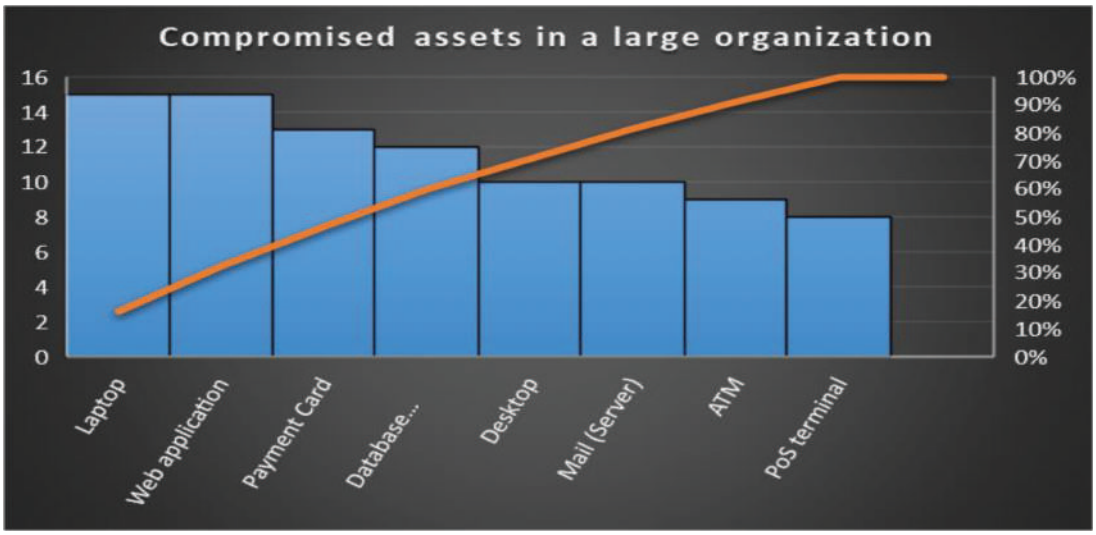

FIGURE 20 | Compromised assets in a large organization.

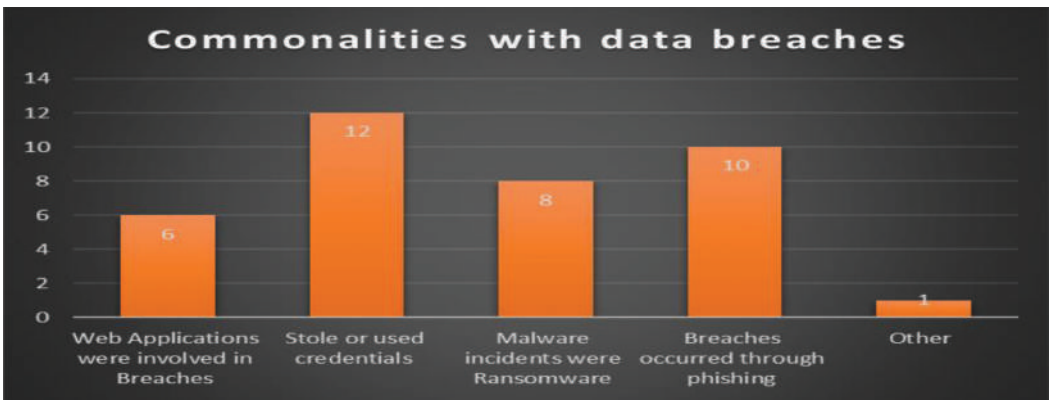

FIGURE 21 | Commonalities with data breaches.

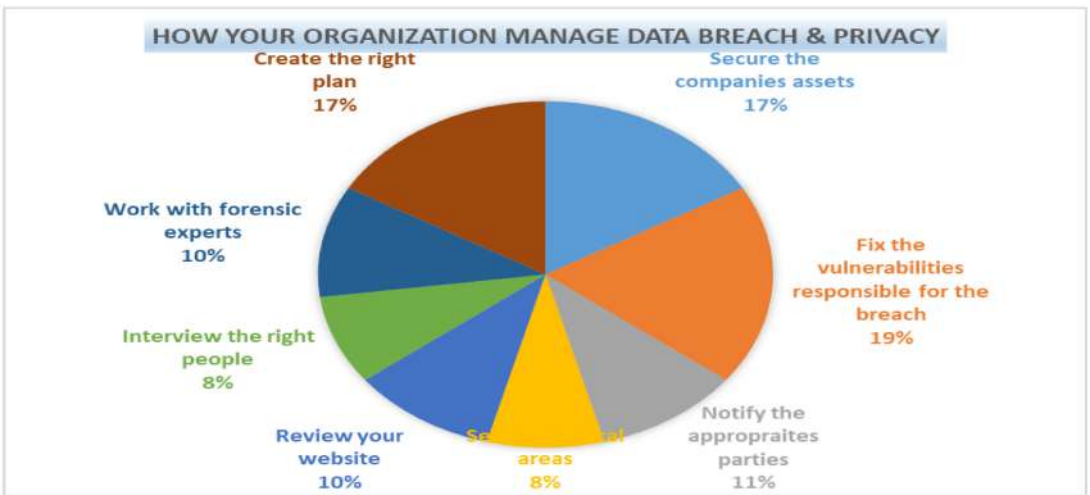

FIGURE 22 | How your organization manages data breaches and privacy.

endpoint devices are secured if the cybersecurity program considers the entire ecosystem without exception. If a single asset is vulnerable, it becomes the last path for cyberthreat to exploit to make its way into the organization's network. Each endpoint's vulnerability status within the enterprise is the only way to compromise the asset.

The respondents share their views, and this is graded in percentage. As many as 95% of respondents categorized web applications and laptop assets as top to be compromised by cyberthreats. A high percentage of data breach incidents take place through these two assets. However, 85 and 80% of respondents mentioned payment cards and databases as assets compromised in large enterprises, resulting in a data breach. These assets store individuals' PII, and their compromise will result in reputational damage to the organization and financial losses and eventually, bankruptcy. As many as 60% of respondents agreed desktop and mail services constitute assets compromised in a large organization. While 55 and 50% of respondents classified automated teller machine and PoS terminals as critical assets, which can bring about high financial loss, litigation, and reputational damage when compromised.

In some cases, the enterprise doesn't know what they have as assets within their ecosystems. When these assets are compromised, they struggle to identify the endpoint through the compromised network. However, because of the organization's size and the thousands of endpoint devices connected, this will take longer to uncover. Implementing an asset management solution is paramount for every organization. This will provide insight into the security team to classify their assets into tiers for scoping when deploying cybersecurity controls (**Figure 20**).

Data breaches occur through different tactics cybercriminals adopt, to infiltrate different enterprise networks, bringing about data breaches. The cyberthreat vector used is multiple. Their success depends on each endpoint's vulnerable state's components. Ultimately, the maturity and seriousness with which businesses handle matters relating to the security of their environment is what matters. There are several commonalities with data breaches resulting in the financial loss of millions of dollars. For example, in J.P. Morgan in 2014, 76 million records were compromised, and Yahoo had one billion records compromised in 2016. These records are vital information for both individuals and the organization. Twelve respondents stated that stolen or used credentials were commonalities with data breaches. When credentials are compromised, cybercriminals leverage them to launch attacks on that organization. Ten respondents mentioned breaches that occurred through phishing. This is common through social engineering tactics targeting private people and organizations. But the victims are mainly individuals. Phishing is a persistent threat of cyberattacks and involves criminals sending messages to masquerade as legitimate entities, targeting many organizations daily. The messages sent out direct their victims to a sham website that captures their personal information unknowingly to their target. Eight and six respondents agreed that malware incidents where ransomware and web applications were involved in data breaches are significant commonalities of data breaches (**Figure 21**).

Most companies fail to adequately safeguard themselves, making their businesses vulnerable to data breaches. In these circumstances, it is advisable not to panic but assemble critical stakeholders to develop the methodology based on the enterprise size and the data breach context. The process often requires a forensic team, legal counsel, human resources, communications, investor, and business relations manager, managing external stakeholders. A simplifying approach to mitigating data breaches addresses how they happen, how they are being rectified, and what those impacted need to secure themselves against further damage. However, 17% of respondents state that they secure the company's assets and create the right plan as the enterprise's top priority to manage unfortunate incidents from data breaches. Further, 19% of respondents mentioned fixing the vulnerabilities responsible for the breach. A cyber incident is due to vulnerability; when the incident occurs and the root cause is identified, the following line of action is to fix the vulnerability. As many as 11 and 10% of respondents' responses explicitly refer to the following steps: notifying appropriate parties, engaging a forensic expert, and assessing the official company website. Few respondents state an interview of the right people and secure physical areas are additional steps to uncover the actors behind the data breach incident and it increases security (**Figure 22**).

Conclusion

As data breaches repeatedly inflict harm on organizations globally regardless of their size, business owners and governments are working towards consolidating existing requirements to curb the increasing data breaches fully. Statistics have shown that the number of data breaches and records globally and cyberattacks is increasing. This number is exponentially increasing in an upward trend. Different countries have their records of data breaches, and the summation of the incident is terrifying, and it creates panic. When the total financial losses resulting from data breaches are quantified in terms of the dollar amount, it will be more than a billion dollars annually.

Furthermore, findings from the research and analysis of the respondents' feedback establish that data breaches crime should be taken seriously by business owners, individuals, and governments. Attackers are more sophisticated and consistently acquiring better technology to facilitate seamless breakthroughs into an enterprise network to launch harm and make financial gains from that effort. Although each country has promulgated laws to end crime, both individuals and organizations suspected of violating the law should face litigation, which may end up in a fine and serve several years in prison for the degree of the crime.

Moreover, non-compliance with regulatory requirements is high across various sectors. They are making the operating environment suspicious of malicious cybercrimes. Only a

few organizations have a data protection plan or strategy supported by technology tools like file integrity monitoring (FIM). This tool allows the enterprise to automate monitoring of its essential files, systems, network, and more.

However, with adequate security investment, the organization should acquire the proper file integrity monitoring software to consistently monitor and detect suspicious real-time changes in their infrastructure ecosystem. Acquiring this tool starts with developing a strategy that clearly defines roles, the correct documentation, and well-thought-out planning. Acquiring a technology tool will increase data security and expedite compliance with cybersecurity standards like PCI-DSS and regulatory agencies like HIPAA. In addition, conforming to both NIST and ISO standards is important to highly secured organizations' ecosystems against the infiltration of hackers.

Additionally, organizations entrusted with customers' data need an excellent cybersecurity program holistically implemented without exception. This is because data breaches tarnish the organization's reputation when they occur. Though hackers are equally enhancing their tactics in acquiring sophisticated technology to easily compromise the networks of their target victims, modern cybersecurity technology such as File Integrity Monitoring (FIM) and Multi-Factor Authentication (MFA) with a robust data protection policy safeguards the enterprise, customers' data, and the organization's reputation.

References

1. IBM Corporation. *Cost of a Data Breach Report 2020*. Armonk, NY: IBM Corporation (2020).
2. Adman S. *Sophos News*. Mumbai: Sophos News (2021).
3. Sarsfield S. *The Data Governance Imperative*. London: IT Governance Publishing (2009).
4. Eryurek E, Gilad U, Lakshmanan V, Kibunguchy-Grant A, Ashdown J. *Data Governance: The Definitive Guide*. Sebastopol, CA: O'Reilly Media, Inc (2021).
5. Castanedo F. *Understanding Data Governance*. Sebastopol, CA: O'Reilly Media, Inc (2017).
6. Petrella A. *What Is Data Governance? Understanding the Business Impact*. Sebastopol, CA: O'Reilly Media, Inc (2020).
7. Arbanas J, DeJong C, Aga G, Sutter D. *Treat your Data Governance Program like the Superpower it is: Making a Case for Data Governance for TMT Companies*. London: Deloitte (2019).
8. Eckerson W. *Creating an Enterprise Data Strategy: Managing Data as a Corporate Asset*. New York, NY: McGraw-Hill (2011).
9. MacDonald A. *Building a Winning Data Strategy*. Cambridge, MA: MIT SLOAN (2020).
10. Tudor J. *Empowering a Data Culture from the Inside Out*. Cambridge, MA: MIT (2020).