Science and Technology 2023, 13(2): 29-35
DOI: 10.5923/j.scit.20231302.04

Impact of Controls on Data Integrity and Information Systems

Sasidhar Duggineni

Compliance Manager at PPD part of Thermo Fisher Scientific, USA

Abstract Data has changed and revolutionized the world around us. With the amount of data use and its need increasing every day, there comes an essential responsibility and a requirement of maintaining the integrity and security of the data. Data integrity and data security are critical issues in today's digital and electronic world, as organizations rely increasingly on electronic data intake, storage, and transmission. Data integrity refers to the accuracy, reliability, and consistency of data over its lifecycle, while data security usually refers to the protection of data from unauthorized access or modification or deletion or theft. This research paper aims to explore the various measures and new improvisation techniques that organizations can take to enhance and retrospect the existing controls for ensuring the integrity and security of their data in accordance with their business, applicable legal and regulatory requirements. These measures may include but not limited to data encryption, access control, data backup and recovery, audit trails, data privacy, cybersecurity, legal and regulatory compliance. This research paper will discuss the benefits and challenges for implementing these measures, as well as best practices for achieving highest level of data integrity and security. Ultimately, the goal of this research is to provide organizations with a comprehensive control option for data integrity and delve into new improvisation controls and techniques in product development to pro-actively address data element risks in a more effective manner. We will also further look at the different ways data can be managed and secured by having adequate controls for data integrity, security, and confidentiality.

Keywords Data Governance, Data Integrity, Regulatory Compliance, Data Management, Clinical Data Integrity

1. Background

In the digital age, data has become an important asset for organizations of all sizes and industries. From financial records, customer information, intellectual property and proprietary research, data is essential for the operation and success of modern businesses. However, as organizations rely more heavily on electronic data storage and transmission, the risk of data loss or compromise increases exponentially. Hence, data integrity and data security evolved to become an integral aspect of data life cycle or system development life cycle (SDLC).

Data integrity is elucidated as the accuracy, reliability, and consistency of data over its lifecycle. In simpler terms, data integrity ensures that data is not corrupted or modified in an unauthorized manner, either intentionally or unintentionally. These concepts are important in the context of data storage, data flow and data management because they ensure that data is protected from unauthorized access and modifications, and that it is accurate, untampered, and reliable. On the other hand, data security, refers to the protection of data from security exploits. It involves measures to prevent unauthorized parties from accessing, stealing, or altering data, as well as ensuring the confidentiality and privacy of sensitive information. Ensuring data integrity and security is crucial for the success and reputation of any organization. Data breaches can have serious consequences, including financial losses, damage to reputation, and legal liabilities. Therefore, it is essential for organizations to have robust measures in place to safeguard their data. In the context of data storage, data flow and data management, data security measures could include encryption, access controls, backup, and recovery procedures, while data integrity measures could include but not limited to checksums, version control, audit trails and log monitoring. Both data security and data integrity are essential for reliable and effective operation of systems and business critical processes that rely on data. Figure 1 and Figure 2 presents the data for Global number of data breaches with confirmed data loss from November 2020 to October 2021, by organization size and target industry respectively [4].

Data breaches occurred at 5212 companies across the globe between November 2020 and October 2021. Financial organisations saw the largest amount of data compromises among the list of industries mentioned above. In terms of organisation size, smaller businesses were more frequently affected by data breaches than major corporations.

* Corresponding author:
Sasidhar.duggineni@gmail.com (Sasidhar Duggineni)
Received: May 12, 2023; Accepted: May 31, 2023; Published: Jun. 2, 2023
Published online at http://journal.sapub.org/scit

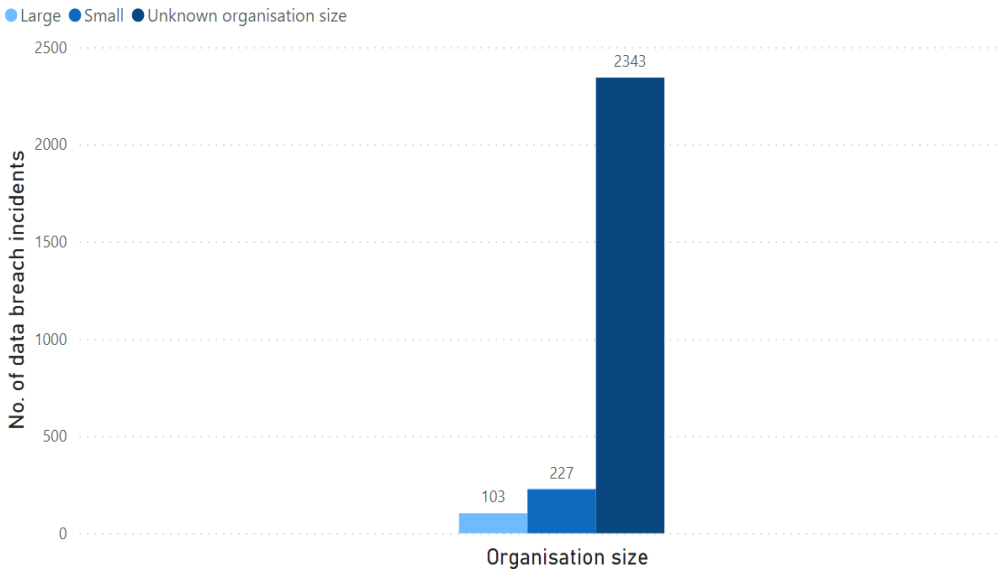

Figure 1. No. of data breach incidents by organization size (November 2020 to October 2021)

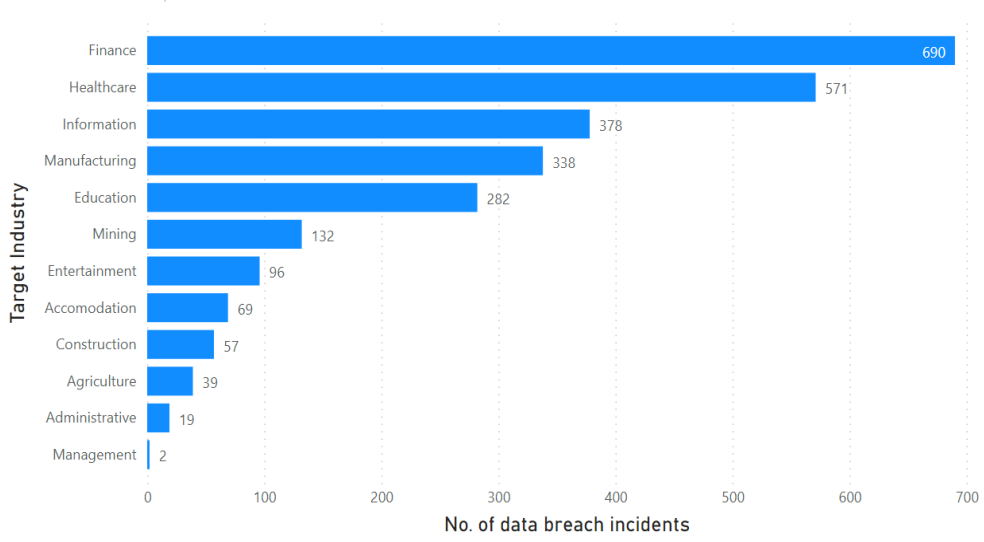

Figure 2. No. of data breach incidents by industry (November 2020 to October 2021)

2. Regulatory Implications for Data Integrity and Data Security

Meeting Regulatory Expectations for Data Integrity is crucial for many local, national, and global organizations. These expectations are enforced in the form of regulations, rules, and laws. In United States, regulations meant to ensure data integrity in pharmaceutical industry are found in several parts of 21 Code of Federal Regulations and have been enforced by the US Food and Drug Administration Agency for decades. 21 CFR Part 11 went into effect in 1997 to extend data integrity regulations into the modern era with electronic records and electronic signatures. World Health Organization, European Medical Agency, and United Kingdom's Medicines and Healthcare Products Regulatory Agency has introduced their own guidance's for data integrity compliance. Sarbanes-Oxley Act (SOX) is a US federal law which sets strict standards for safeguarding integrity of financial data. United States Health Insurance Portability and Accountability Act (HIPAA) provides federal protection for patient's health data against misuse or exposure and requires technical and administrative controls to ensure compliance with HIPAA.

3. Methods (An Experimental Case Study)

There are several things to consider when evaluating the dangers to data security and data integrity in reference to any given system. This is explained with the help of an example, an ecommerce application called Order Management System (OMS) which entered development in the year 2018 to intake, store and manage sensitive customer data for a large retail company. The main business use case for the application is

to track customer purchases, preferences and to provide personalized recommendations and offers.

During the requirements gathering and analysis phase, the data security requirements were not adequately defined, analysed, and documented. Further, to trace the consequential data breach incidents the integrity measures such as providing administered access controls and maintaining audit trails were missing. The development team did not receive adequate clearly defined data integrity requirements from the business stakeholders and security standards during the code building phase did not adequately assess the potential risks to the customer data. This lack of proper requirements gathering, and analysis led to a design that did not consider the best practices of data security thoroughly. As a result, the data security and data integrity requirements were incomplete and did not adequately address the security controls needed for handling of the customer data. In addition, the application did not include adequate authentication and version controls to protect the customer data. During the coding phase, the developers left off important secure coding practices like encrypting data with modern cryptographic algorithms, password management, default deny, input validation and output encoding.

Furthermore, the application was not tested comprehensively for data security vulnerabilities before being deployed to production as the testers and other stakeholders were not adequately trained on data security best practices. As a result, several vulnerabilities went undetected, proper security protocols were not followed while working with the customer data and potential vulnerabilities were not reported in a timely manner.

When the application was deployed, it was discovered that the systems and infrastructure were not properly configured to secure the data. The customer account and financial data was not encrypted, and the insecure coding methods implemented by the developers made the platform powerless against the Packet Sniffing attacks during the data flow from the frontend to backend. The network was not properly segmented, and the database backups did not have limited access, making it easy for attackers to access the customer data. Consequently, there were multiple data breaches by the hackers over a period for which the stats can be viewed below in the Table 1. As the system was vulnerable to attacks by the hackers, it became an easy target once they could peak into the system and multiple attacks took place which led to breach of millions of user records which is described in the fig. 4 below.

4. Results

Ergo, the company faced legal action and financial penalties, as well as a loss of customer trust and loyalty.

Fig 4 represents that type of records that were stolen in the data breach incidents with their total number and the corresponding share to total stolen records. To prevent similar issues in the future, the retail company could implement several changes to its SDLC process which will be discussed in the upcoming section on the paper.

Table 1. Yearly no. of data breach incidents

Year	No. of data integrity / security incidents
2018	4
2019	9
2020	6
2021	3

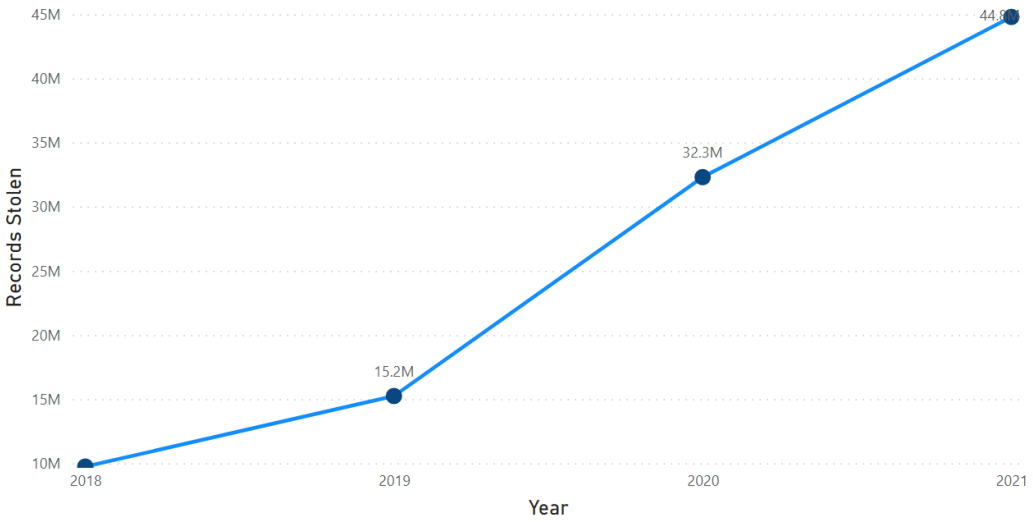

Figure 3. No. of records stolen and corrupted yearly

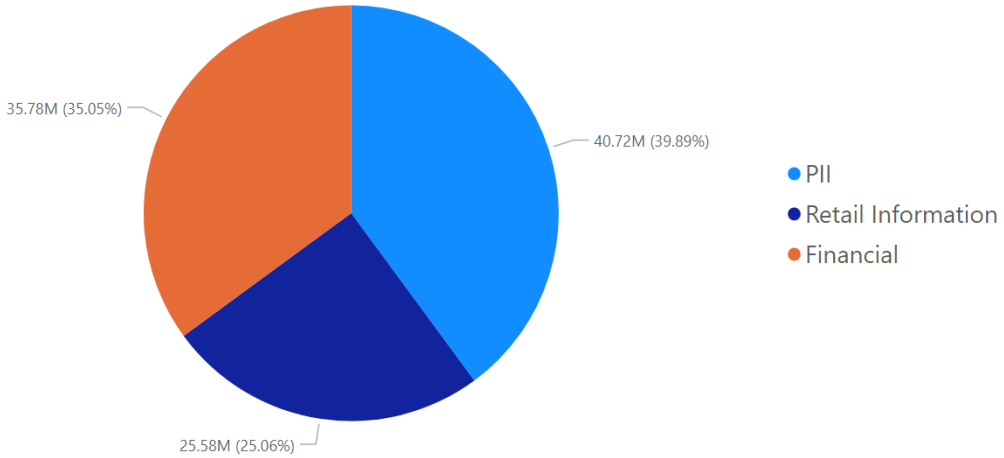

Figure 4. No. of data records stolen for various data categories

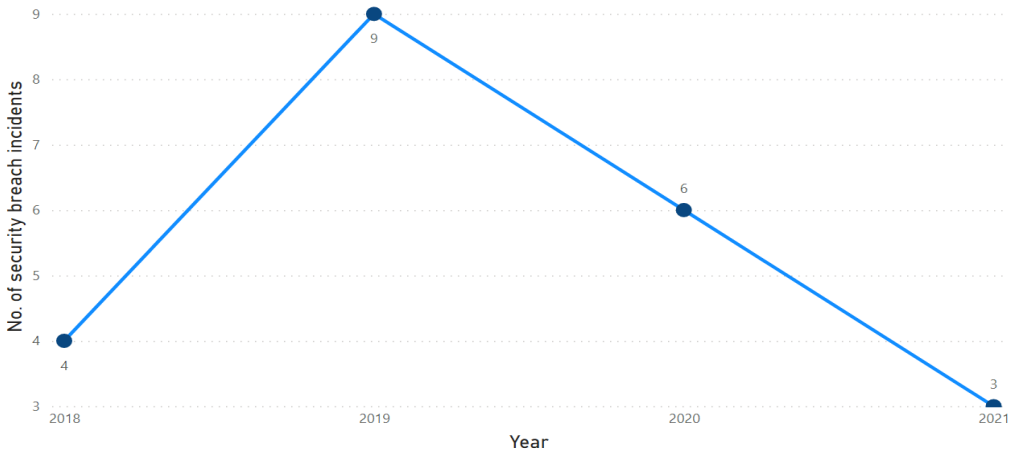

Figure 5. No. of data breach incidents year wise

Other types of attacks that have compromised the data security and integrity include:

- SQL injection attacks: These attacks involve injecting malicious code into a database through a website or application form, to access or manipulate data.
- Phishing attacks: These attacks involve sending fraudulent emails or text messages that appear to be from a legitimate source, to trick people into divulging sensitive information or clicking on malicious links.
- Ransomware attacks: These attacks involve using malware to encrypt data and hold it hostage until a ransom is paid to the attackers.

Now we'll study another exposition of an application called "Electronic Payment Management System" (EPMS) to understand the importance of having a secured systems with reference to data security and integrity. This new age Hospital patient billing system which entered development in the year 2018 to intake to automate the manual task of billing to ensure faster payment, improve workflow, and tracking patient information.

The coding standards, business requirements, user requirements were all well incorporated with required data security and data integrity specifications, and as a result this application was developed using a robust and secure coding practices in software development life cycle with secure infrastructure. The development team further consulted with security experts to assess the potential risks to the Personal Identifiable Information (PII), health care data and financial data. As a result, the data security and data integrity requirements were comprehensive and properly addressed the security and integrity of the data. The data storage platform architecture to store the data was protected by multiple layers of security and multi factor authentications. In addition, the application encompassed high level regulation and access controls to protect the financial and personal data. In the build phase, the developers used secure industry standard coding practices and properly validated user input, which minimized vulnerabilities in the application and properly handled sensitive data such as passwords and personal information of the users by keeping it encrypted. The system integration and user acceptance testing were thoroughly carried out for data security vulnerabilities, data integrity controls and any vulnerabilities that were discovered were promptly fixed before the deployment. Post the deployment stage, the systems and infrastructure were properly configured to secure the data.

Hence, making the unauthorized access difficult to access the data.

As a result of these measures, the application "Electronic Payment Management System" (EPMS) encountered around 90% less data security breaches compared to Order Management System (OMS) which can be studied through the comparison of table 2 and table 1. The financial data was protected from external threats and the bank did not experience any major data breaches or other security incidents. The organization was able to maintain the trust and loyalty of its customers and avoided any legal or financial consequences. Fig. 6 briefly describe the number of data security breach incidents on the Electronic Payment Management System.

Table 2. Yearly no. of data breach incidents

Year	No. of data integrity/security incidents
2018	1
2019	1
2020	3
2021	0

The table 3 and fig. 7 below summarises the comparison of the two systems based on the functional and regulatory compliance requirements for data security and integrity.

Table 3. Compliance requirement support

Compliance Requirement	EPMS Application	OMS Application
Privacy	Fully Supports	Partially Supports
Audit Trails	Fully Supports	Partially Supports
Data Storage Protections	Fully Supports	Partially Supports
Audit Readiness	Fully Supports	Doesn't Support
Regulatory Compliance	Fully Supports	Partially Supports
Legal Compliance	Fully Supports	Doesn't Support
Digital Forensics suitability	Fully Supports	Partially Supports
Data Confidentiality	Fully Supports	Partially Supports
Retrievability and accessibility	Fully Supports	Doesn't Support
Traceability	Fully Supports	Partially Supports
Reliability	Fully Supports	Partially Supports

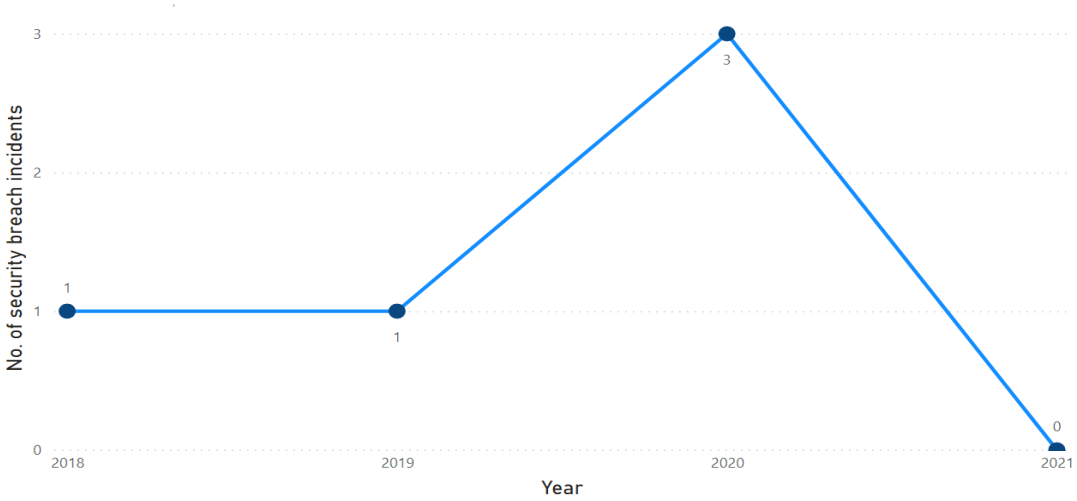

Figure 6. No. of data breach incidents year wise

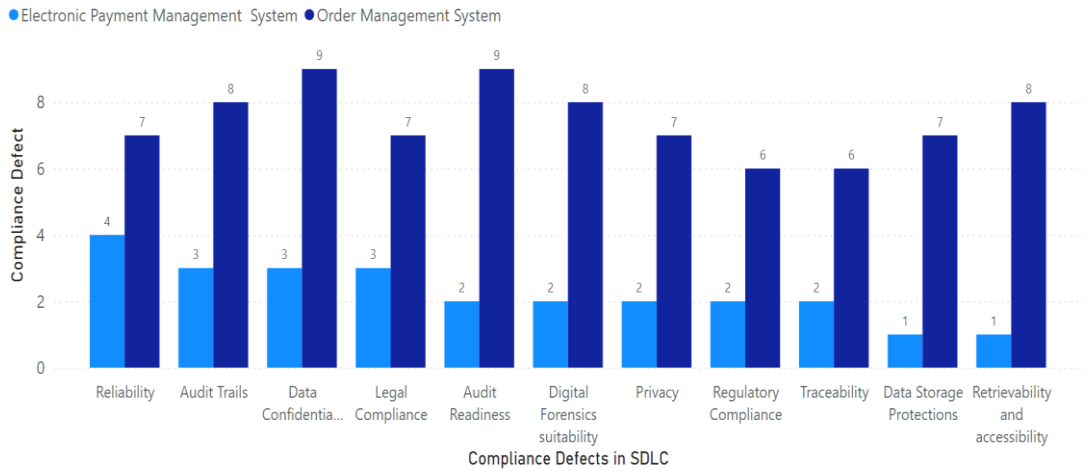

Figure 7. Quantitative Analysis of Compliance State for both systems (EPMS in Light Blue and OMS in Dark Blue)

5. Data Security and Data Integrity Measures

By the above elaborated case studies, we have understood the significance of having systems which are robust in terms of securing the data and maintaining its integrity. And to build such systems there is a need to keep an eye on certain procedures which include conducting thorough assessment of data security risks at the beginning of the project, using secure coding practices, and testing techniques, properly configuring systems and infrastructure, and training developers and other stakeholders on best practices for assuring data integrity and data security.

Below listed are some of the highly impactful measures for ensuring the integrity and security of data throughout its life cycle to keep it complete, accurate and safe.

- **Validating input data**

Before entering the servers or data storage system or database, input data should always be verified. It is the process of ensuring that data is accurate and trustable. Regardless of the source of the data, whether it is being ingested from internal systems, external sources, or end-users of an application, the data should be validated for accuracy and trustability.

- **Implementing access controls**

Data access permissions should be strictly controlled to guarantee that only individuals with the necessary authorizations have access to the data. Access should only be allowed to those who require it and adhere to a least privileged model of security. Data should be segregated to reduce the likelihood of unwanted access within different groups of users.

- **Keeping audit trails**

Maintaining an automatic audit monitoring system that can identify the origin of data changes is essential. In the event of a data breach, the audit trail should also be able to identify the breach's origin and keep track of data events like creation, deletion, and updating as well as the moment they took place and the person who performed them.

- **Backing up data**

A backup procedure makes sure that no data is lost in the event of a data loss. Hence, it is crucial to have regular, trustworthy, and timely backups of data systems which have controlled access to limited users.

- **Guiding the staff**

The employees in an organization should be educated to always maintain the integrity of data in all work processes. It is important to build a culture of effective data management where team members are always encouraged to handle data in a way that assures the consistency and dependability of data and where individuals abide by the principles of data integrity.

- **Code development improvisations**

System development can be reinforced by incorporating a variety of controls during various phases of system development. There is no one size fits all approach to accomplish this goal. Starting from the coding standards developers can be proactive and incorporate data integrity and data quality as a code flowing through the build and deploy processes. Examples of these controls include but not limited to ensuring code doesn't carry integrity violations such as modules/libraries from untrusted sources, performing code reviews with targeted data integrity benchmarks, development of critical system functionalities with required attributes such as audit trails, traceability, authentication, enforcing integrity of values in database columns and rows, implementing logs and input data validations. During the stages of requirement gathering for the product, emphasis should be given on effective translation of data integrity requirements into formal product requirement documents for easy conveyance of data integrity expectations to product designers/developers from business user community.

6. Conclusions

Throughout the course of this article, we have deeply investigated the possible causes of data integrity, data security breaches and the technical and procedural measures that can be followed and improvised to keep the user data confidential, accurate and reliable for businesses. By putting data security and integrity measures into place, a business can protect itself from data breaches, unnecessary financial expenditures, loss of public trust, potential threats to brand reputation, and loss of future revenue. A coordinated data security architecture also protects the organisation from the negative legal and regulatory repercussions associated with data breaches and guards against unwanted access to computers, laptops, smart devices, websites, networks, and devices.

Hence, it is important for individuals and organizations to be fully cognizant of these ever-evolving types of threats and to take new improvised steps to better protect themselves and their data against them. All things considered; data security and integrity are a multidisciplinary team effort that should be tackled on a comprehensive scale.

Competing Interests

Author reports no Competing Interests.

REFERENCES

[1] M. A. Shah, R. Swaminathan, and M. Baker, "Privacy-preserving audit and extraction of digital contents," IACR Cryptology EPrint Archive, vol. 186, 2008.

[2] D. Ferraiolo, R. Sandhu, S. Gavrila, R. Kuhn and R. Chandramouli, "Proposed NIST Standard for Role-based Access Control", ACM Trans. Information and System Security, vol. 4, no. 3, pp. 224-274, 2001.

[3] E.B. Fernandez, R.C. Summers and C. Wood, Database Security and Integrity., Feb. 1981.

[4] https://www.statista.com/statistics/329608/security-incidents-confirmed-data-loss-industry-size/#:~:text=Data%20breaches%20worldwide%202020%2D2021%2C%20by%20target%20industry%20and%20organiyation%20size&text=Between%20November%202020%20and%20October,highest%20number%20of%20data%20violations.

[5] E. Ferrari and B.M. Thuraisingham, "Security and Privacy for Web Databases and Services", Advances in Database TechnologyEDBT 2004 Proc. Ninth Int'l Conf. Extending Database Technology, Mar. 2004.

[6] E. Bertino, E. Ferrari and L. ParasilitiProvenza, "Signature and Access Control Policies", Proc. 2003 European Symp. Research in Computer Security (ESORICS-03), Oct. 2003.

[7] B. Thuraisingham, "Mandatory Security in Object-Oriented Database Systems", Proc. Int'l Conf. Object-Oriented Programming Systems Languages and Applications (OOPSLA), 1989.

[8] R. Agrawal, R. Srikant and Y. Xu, "Database Technologies for Electronic Commerce", Proc. Very Large Databases Conf. (VLDB), 2002.

[9] Trusted Computer System Evaluation Criteria, 1975.

[10] Deyan Chen and Hong Zhao, "Data Security and Privacy Protection Issues in Cloud Computing", 2012 IEEE International Conference on Computer and Electronics engineering.

[11] Mohamed Al Morsy, John Grundy, Ingo Müller, "An Analysis of The Cloud Computing Security Problem," in Proceedings of APSEC 2010 Cloud Workshop, Sydney, Australia, 30th Nov 2010.

[12] S. Subashini, V. Kavitha. A survey on security issues in service delivery models of cloud computing. Journal of Network and Computer Applications 34(2011)1-11.

[13] M. Lori, "Data security in the world of cloud computing," Co-published by the IEEE Computer And reliability Societies, pp. 61–64, 2009.

[14] C. Wang, K. Ren, W. Lou, and J. Li, "Toward publicly auditable secure cloud data storage services," Network, IEEE, vol. 24, no. 4, pp. 19–24, 2010.

[15] L. Wei, H. Zhu, Z. Cao, X. Dong, W. Jia, Y. Chen, and A. V. Vasilakos, "Security and privacy for storage and computation in cloud computing," Information Sciences, vol. 258, pp. 371–386, 2014.

[16] C. Wang, Q. Wang, K. Ren, and W. Lou, "Privacy-preserving public auditing for data storage security in cloud computing," in INFOCOM, 2010 Proceedings IEEE. Ieee, 2010, pp. 1–9.

[17] Etzioni. Medical records --- enhancing privacy. preserving the common good. Hastings Center Report, 23(2): 14--23, 1999.

[18] F. A. Lategan and M. S. Olivier. Enforcing privacy by withholding private information. In S. Qing and J. H. P. Eloff, editors, Information Security for Global Information Infrastructures, pages 421--430. Kluwer, 2000.

[19] J. Garret. John rawls on moral principles for individuals: With emphasis on implications for business ethics, February 2002.

[20] Cong Wang, Qian Wang, Kui Ren, Ning Cao and Wenjing Lou, "Towards Secure and Dependable Storage Services in Cloud Computing", IEEE Transactions on Services Computing, vol. 5, no. 2, pp. 220-232, May 2011.

[21] Breached Patient Records Tripled in 2018 vs 2017 as Health Data Security Challenges Worsen, Oct. 2018, [online] Available: https://www.protenus.com/press/press-release/breached-patient-records-tripled-in-2018-vs-2017-as-health-data-security-challenges-worsen.

[22] Healthcare Data Breaches Reach Record High in April, Oct. 2019, [online] Available: https://www.modernhealthcare.com/cybersecurity/healthcare-data-breaches-reach-record-high-april.

[23] T. D. Oyetoyan, M. G. Jaatun and D. S. Cruzes, "Measuring developers' software security skills usage and training needs" in Exploring Security in Software Architecture and Design, Hershey, PA, USA: IGI Global, vol. 1, 2019.

[24] G. Manogaran, C. Thota, D. Lopez and R. Sundarasekar, "Big data security intelligence for healthcare industry 4.0", Cybersecur. Ind., vol. 4, pp. 103-126, Apr. 2017.

[25] J. Bethencourt, A. Sahai, and B. Waters, "Ciphertext-policy attribute-based encryption," in Security and Privacy, 2007. SP'07. IEEE Symposium on. IEEE, 2007, pp. 321–334.

[26] C. Erway, A. Kupc¸ ¨ u, C. Papamanthou, and R. Tamassia, "Dynamic provable ¨ data possession," in Proceedings of the 16th ACM conference on Computer and communications security. Acm, 2009, pp. 213–222.