International Journal of Current Engineering and Technology
E-ISSN 2277 – 4106, P-ISSN 2347 – 5161

Available at http://inpressco.com/category/ijcet

Research Article

Compliance-Driven Data Governance: A Survey on GDPR, and HIPAA in Cloud Databases

Anirudh Parupalli[1*] and Suhag Pandya[2]

[1]Independent Researchers

Received 01 Dec 2022, Accepted 20 Dec 2022, Available online 21 Dec 2022, **Vol.12, No.6 (Nov/Dec 2022)**

Abstract

Data management by allowing organizations to manage huge volumes of sensitive data in distributed cloud databases. Although this paradigm provides scalability and operational efficiency, it is in violation of legal frameworks, it has significant privacy and security problems, particularly with The Health Insurance Portability and Accountability Act (HIPAA) and the General Data Protection Regulation (GDPR). Data governance has become a crucial field in compliance that focuses on the legality and compliance of a given data practice, but also the effectiveness and reliability of cloud data. This article provides an exhaustive survey of data governance of cloud databases where specific focus on GDPR and HIPAA. It looks at the main principles of data stewardship, classification and retention and deletion policies, anonymization, encryption, and auditability and examines how they apply in cloud settings. Moreover, new methods, such as those based on AI with regard to compliance frameworks and metadata-based governance, are considered to outline their benefits and drawbacks. The results reveal that there are still ongoing challenges of balancing rigorous governance processes that can provide privacy, security and compliance in dynamic distributed cloud systems. The study offers a systematic insight to guide practitioners and researchers to come up with effective, scalable, and compliant data governance models that find application in cloud-based environments.

Keywords: *Cloud Computing, Data Governance, GDPR, HIPAA, Cloud Databases, Compliance, Privacy, Security.*

Introduction

Cloud computing's scalability, cost-effectiveness, and on-demand computer resources have transformed the way companies handle, store, and evaluate data [1]. Another important element of cloud infrastructure is cloud databases that allow storage and retrieval of huge amounts of structured and unstructured information in distributed settings [2][3]. This paradigm shift has fueled the digital transformation in all sectors, such as e-commerce, healthcare, and finance, where delicate and mission-critical data are being increasingly outsourced to the cloud, and the use of cloud databases is of great concern in terms of secrecy, data privacy, and regulatory compliance [4][5], especially when it comes to private data such as personally identifiable information (PII) and protected health information (PHI).

Strict regulatory requirements, such the General Data Protection Regulation (GDPR) in the EU and the Health Insurance Portability and Accountability Act (HIPAA) in the US, have created systems to monitor and protect sensitive data [6][7].

*Corresponding author's ORCID ID: 0000-0000-0000-0000
DOI: https://doi.org/10.14741/ijcet/v.12.6.18

HIPAA requires health data to be available, accurate, and private in the healthcare industry, but GDPR places more emphasis on data subjects' rights, accountability, and transparency as well as the legitimate handling of personal data [8][9][10]. Such regulations also place an obligation on data controllers and processors as well as the cloud service providers (CSPs) which have an obligation to ensure that their infrastructure and service is compliant. Therefore, there is a need to insist on effective governance processes in organizations using cloud databases in order to comply with the regulations, to prevent fines, and to ensure trust in the stakeholders [11].

Data Governance in the cloud entails data policy, data roles, data processes and data controls that take into account the correct handling of the data assets by covering all aspects of its life cycle. A cloud scenario is, however, where governance is complicated to an extent that it is affected by the multi-tenancy, data residency problems, vendor lock-in and the shared responsibility paradigm of the CSP and the customer [12][13][14]. To support GDPR and HIPAA in the distributed and dynamic environment, the organizations must have clear data ownership, robust data security and privacy mechanisms, continuously monitor operations and keep an auditable record of

data processing processes [15][16][17]. Compliance in cloud environments is further complicated by the absence of direct control of infrastructure, and thus effective data governance is required as a support mechanism to GDPR and HIPAA compliance in the case of cloud databases. This paper attempts to bridge this gap through a systematic review of the concepts, issues and approaches of compliance-based data governance [18]. By analyzing current practices, tools, and technologies, the paper highlights how organizations can align their cloud data management processes with regulatory requirements without compromising on scalability and performance. Cloud data governance demands a systematic examination to guide organizations in mitigating legal and operational risks.

Structure of the Paper

This paper is organized as follows: Section II. foundations of compliance-driven data governance. Section III Regulatory frameworks: GDPR and HIPAA Section IV. cloud database adaptations for compliance in Section V Literature review, Section VI Conclusions and future work.

Foundations of Compliance-Driven Data Governance

In the data-driven world of today, businesses create and handle enormous volumes of sensitive and personal data, a large portion of which is stored in dispersed cloud environments. This necessitates organized data governance procedures that guarantee data quality in addition to and usability but also explicitly address regulatory and legal obligations. Laws like the GDPR and the HIPAA state that the primary objective of compliance-driven data governance is data management, which imposes strict requirements on privacy, security, and accountability. Unlike traditional governance, which prioritizes operational and business value, compliance-driven governance ensures that collecting, preserving, processing, and disseminating data are done in a manner that complies with the law, ethics, and transparency.

In cloud computing environments, where data may be stored across jurisdictions and managed by third-party providers, the challenge is even greater. Here, compliance-driven governance frameworks provide mechanisms for Access control, anonymization, encryption, and data classification, and monitoring ensuring that organizations retain oversight and accountability even when leveraging scalable cloud services [19] Through extensive network access, it offers metered service that is available whenever needed. It is crucial to describe the three cloud service architectures shown in Figure 1:

Fig.1 Cloud Service Model

Infrastructure as a Service (IaaS): The infrastructure resources that offer the most control and management are these. At the same time, it takes the greatest customer effort to get the solution into production.

Platform as a Service (PaaS): It provides a platform for developing applications, while cloud providers manage the supporting infrastructure.

Software as a Service (SaaS): In this arrangement, the client gets a ready-made solution. It offers the least amount of customization yet necessitates the least amount of setup and administration work.

The cloud database contains data on unique server farms located in various locations. The goal database management framework is not the same as the cloud database structure. Many hubs are located in different geographical locations and are designed to administer questions for server farms also known as corporate farms through a cloud database. His connection is required for the database on the cloud administrations to be fully accessible [20].

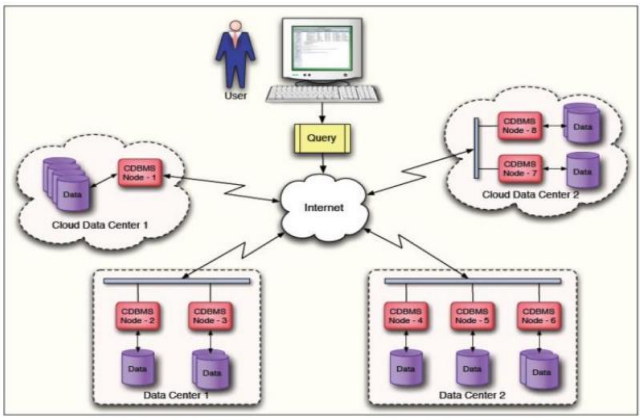

Fig.2 Structure of Cloud Database

In order to profit from databases in the cloud, several systems have been devised (see Figure 2). The user can benefit from it via a mobile device that can access the cloud database using 3G/4G services or a home computer via the internet. The structure of cloud databases is shown in order to comprehend their infrastructure.

Core Principles of Data Governance

A fundamental grasp of the following principles is essential for effective data governance in AI-powered business analytics ecosystems: data lineage, stewardship, metadata management, and data quality. The technique of following data from its origin through modifications, known as data lineage, enables accountability and transparency [21]. Data traceability is essential for AI systems because it facilitates model validation, repeatability, and bias reduction.

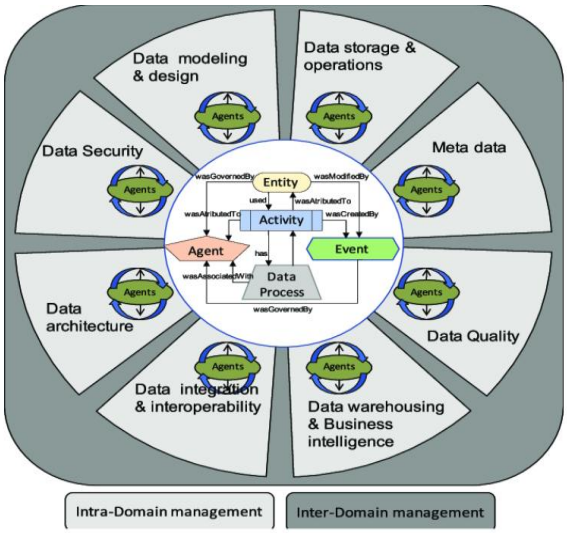

Fig.3 System Architecture for Data Governance

Data governance framework structured around core elements Entity, Activity, Event, Agent, and Data Process which form the central hub for managing and monitoring data flows in Figure 3. Surrounding this core are eight critical governance domains: Data security, architecture, modelling and design, metadata, data warehousing and business intelligence, data integration and interoperability, and data operations and storage. Each domain is managed by specialized agents responsible for enforcing rules, ensuring consistency, and maintaining quality and compliance. The framework also differentiates between intra-domain management (governance within individual domains) and inter-domain management (coordination across domains), enabling both localized control and holistic integration. Such agent-based, multi-domain is robust in data governance, aligned to regulations (e.g., GDPR or HIPAA) and high-quality data operations in complex enterprise environments.

Governance Frameworks and Organizational Policies

Data governance underpinned by compliance is anchored by governance frameworks and organizational policies and offers a well-structured approach to ensuring that all data activities are consistent with the requirements of the law and regulations. An organization's high-level structure, responsibilities, and protocols that regulate data management are described in a governance framework. It guarantees the fact that the compliance is not referred to as an ad hoc performance but rather, it becomes a part of the organizational culture and performance in Figure 4.

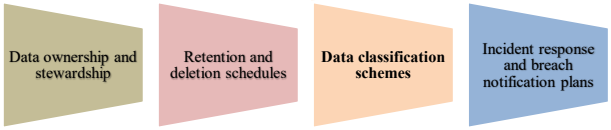

Fig.4 Governance Framework and Organizational Policies

Data Ownership and Stewardship

Data ownership and stewardship define the responsibilities and duties of individuals or teams within an organization who are responsible for handling data in a way that is safe, compliant, and ethical. In compliance-based data governance, data ownership and stewardship offer accountability throughout the data lifecycle, which is essential for laws like HIPAA and GDPR [22]. They provide quality assurance, track daily data compliance with governance guidelines, manage data integrity, and serve as a liaison between business users, IT teams, and compliance officers.

Data Classification Schemes

Data classification schemes refer to a structured method of classifying data according to sensitivity of the data, criticality of the same and regulatory requirements. Under compliance-driven governance, the classification process is a preliminary stage that allows organizations to adopt suitable protection and treatment to various forms of information, and in line with principles, such as data minimization and confidentiality.

Public: The firm uses public data, which is least harmful and least sensitive.

Private: Private data is typically compartmentalized information that has to be kept secret for various reasons but may not harm the business. One type of data that falls within the private category is information on human resources.

Confidential: Internal corporate data that may be less regulated but might be harmful if disclosed.

Sensitive: Data with the highest level of integrity requirements and the least amount of access. This is usually the information that harms the company the most.

Proprietary: Data that is limited in its disclosure to third parties or that includes information that might lower the company's competitiveness is known as proprietary data.

Retention and Deletion Schedules

A retention schedule defines specific timeframes for which different categories of data—such financial transactions, client records, or health information must be kept based on legal, contractual, or business requirements Deletion schedules complement retention policies by specifying how and when data should be securely erased, anonymized, or archived after its retention period expires. This includes defining acceptable methods of secure deletion.

By implementing clear and automated retention and deletion schedules, organizations can

Minimize compliance risks associated with over-retention of sensitive data.

Reduce storage costs and improve operational efficiency.

Demonstrate accountability and adherence to privacy and security obligations.

Incident Response and Breach Notification Plans

In order to fulfil regulatory reporting requirements and enable organizations to promptly identify, evaluate, contain, and recover from data security problems, incident response and breach notification strategies are essential components of compliance-driven data governance. Regulations like GDPR and HIPAA explicitly require timely breach notifications and documented response procedures to minimize harm and maintain accountability [23]. An effective incident response plan outlines the methodical procedure for handling security events, which includes determining and categorizing the level of compromise, putting together a response team, containing the danger, looking into the underlying cause, and returning to regular operations. Roles and responsibilities must be clearly assigned, ensuring coordinated action during a crisis.

Regulatory Frameworks: GDPR and HIPAA

Cloud-based data management systems must comply with stringent legal frameworks that regulate how private data is gathered, used, stored, and shared. HIPAA in the US and the GDPR in the EU are two of the most well-known and often used legislation. Both frameworks impose legal obligations on companies and cloud service providers to ensure accountability, privacy, and data security. The essential guidelines and regulations of HIPAA and GDPR with regard to cloud database governance.

Key Principles of GDPR

A thorough framework for safeguarding personal data is the GDPR that has been in effect since May 2018 and is immediately applicable in all EU Member States [24]. Businesses collect, handle, keep, and disseminate personal data based on 7 fundamental principles, which are depicted in Figure 5.

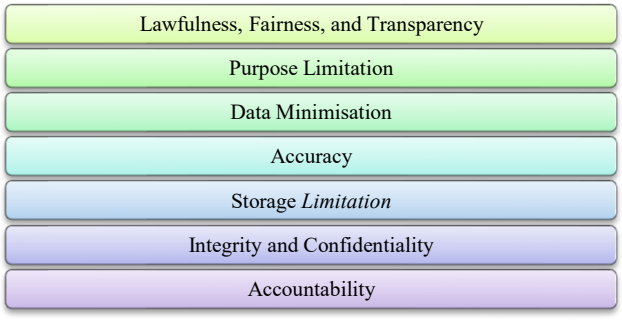

Fig.5 Key Principles of GDPR

Lawfulness, Fairness, and Transparency

The GDPR and the 2018 Act's requirements (especially those stated in Articles 6, 7, 8, and 9) must be met by any controller's processing of personal information to be regarded as legal. Additionally, Personal data cannot be processed or used in any other way without authorization.

The idea of fairness is also somewhat broad, requiring that any processing of personal information be fair to the individual whose information is being processed and steer clear of being very harmful, unexpected, dishonest, or deceptive. One of the main tenets of the GDPR's data protection framework is transparency, which encompasses a variety of obligations and privileges to guarantee that individuals and authorities may access the processing of personal data. Controllers must provide individuals explicit notice regarding the processing of their personal data.

Purpose Limitation

It is crucial that personal information be collected for specified, reasonable, justifiable goals that are established at the time of collecting and that must not be misused that contradicts those goals. Data controllers may handle data for historical, scientific, statistical, or public interest archiving purposes even if they are not thought to be incompatible with the original objectives, as long as adequate safeguards are in place. The 2018 Act's other provisions define when controllers may do extra processing for public interest reasons.

Data Minimisation

Data minimization helps ensure restricting the quantity of data to prevent the loss or theft of data to protect personal information security and integrity in case of a breach. Additionally, it facilitates organizations' efforts to ensure that the personal information they possess is up-to-date and correct, therefore promoting adherence to the principles of accuracy.

Accuracy

According to this concept, controllers are in charge of ensuring that Information about individuals is correct and updated as needed. Controllers are required to do all within their power to guarantee that erroneous personal information is promptly updated or destroyed, taking into account the objectives for processing.

Storage Limitation

Controllers are only permitted to retain personal data for as long as necessary to achieve the objectives for which it uses. In the event that personal information is processed only for historical, scientific, statistical, or

archival reasons in the public interest, the GDPR allows for a longer retention period, provided that the necessary organizational and technical safeguards are in place to safeguard each person's rights and liberties.

Integrity and Confidentiality

To guarantee the proper degree of security and privacy, controllers must only handle personal data in a way that guards against theft, damage, or destruction, as well as processing that isn't permitted or authorized. This requires controllers to use the right organizational or technological tactics.

Accountability

Accountability is increased by following the other data protection principles, which include putting in place a strategy for data protection by default and by design, setting up suitable organizational and technical safeguards, having transparent information that is easy to find, and having clear, explicit data retention guidelines. Additional steps to show adherence to data protection principles include establishing internal rules, adhering to certification programs or codes of conduct, documenting and, if required, disclosing breaches involving establishing appropriate privacy policies and warnings, and protecting personal information.

Key Principles of HIPAA

The US enacted in 1996, the HIPAA creates nationwide guidelines for maintaining the confidentiality and integrity of medical records [25]. HIPAA governs how insurers, business partners, and healthcare providers gather, utilize, disclose, and safeguard Protected Health Information (PHI). The Security Rule, Privacy Rule, and Breach Notification Rule are the two primary statutes that uphold its fundamental tenets. When taken as a whole, these regulations permit the delivery of healthcare services while guaranteeing the privacy, accuracy, and accessibility of medical records (Figure 6).

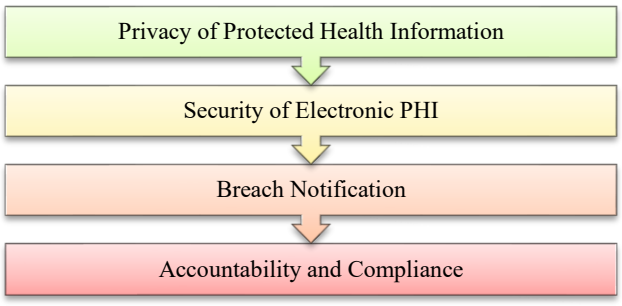

Fig.6 Key Principle of HIPAA

Privacy of Protected Health Information

The HIPAA Security Rule, which is an extension to the Privacy Rule, mandates that organizations have administrative, technological, and physical protections in place to protect the privacy of PHI. Establishments with which PHI usage agreements have been struck. All data that may be used to identify a particular individual is included in PHI [26]. PHI is any part of an individual's medical and billing history that should only be disclosed to those who are authorized.

Security of Electronic PHI

HIPAA provides a thorough system of regulations, including those pertaining to security, privacy, breach reporting, and enforcement, to protect and maintain health information [27]. The necessary security protocols to safeguard electronic protected health information, or ePHI, are described in the security legislation. It consists of administrative, technological, and physical safeguards to maintain the confidentiality, accessibility, and integrity of ePHI. The privacy regulation establishes nationwide guidelines to safeguard people's private health data, including medical records.

Breach Notification

The HIPAA breach notification regulation enables organizations that handle PHI 60 calendar days to notify persons about the breach, its complexity, and the actions taken to reduce current and potential risks. This is to guarantee that the person is aware of the circumstances.

Accountability and Compliance

Accountability and Compliance in data governance describes an organization's capacity to take full responsibility for its data practices to comply with the laws, policies, and ethical principles, and provide evidence of the same. This is done through clarification of obligations and responsibilities, documented policies and procedures, frequent risk assessment, risk auditing, training of staff on data protection requirements, and retention of records (logs, reports) to demonstrate compliance with HIPAA regulations.

Cloud Database Adaptations for Compliance

The traditional data management practice is subject to specific changes to achieve compliance with regulatory frameworks, including GDPR and HIPAA, in cloud databases. The distributed storage, shared infrastructure, and loss of direct control introduced by cloud environments present a set of challenges that require special technical and organizational precautions. In this section, the core adaptations and best practices for businesses to employ while creating their cloud data operations match regulatory guidelines are outlined in Figure 7 data governance framework which is discussed below:

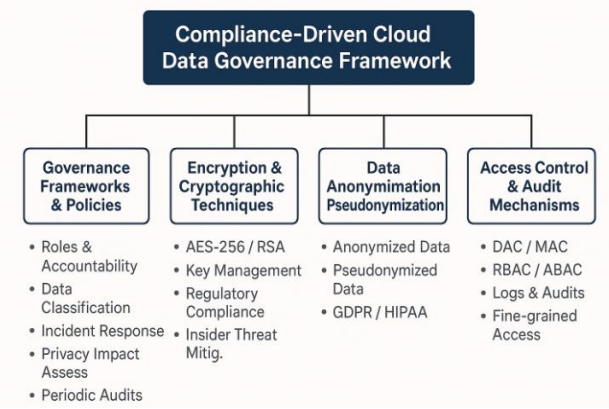

Fig.7 Compliance Driven Cloud Data Governance Framework

Governance Frameworks and Policies

A cloud governance framework is necessary for businesses looking to efficiently manage their cloud resources while mitigating risks and ensuring compliance with compliance-driven data management in cloud databases [28]. They offer an organized set of guidelines, procedures, and obligations to comply with rules like GDPR and HIPAA regulates the gathering, storing, accessing, exchanging, and destroying of data. a strong governance framework defines data ownership and stewardship roles, ensuring accountability for compliance throughout the data lifecycle data is classified distinguishing between sensitive, personal, and public data governance frameworks include incident response procedures, privacy impact assessment processes, and periodic audits to monitor compliance and mitigate risks.

Encryption and Cryptographic Techniques

Data integrity during the transmission and storage procedures. Strong encryption standards are mandated by ordinances like the GDPR and the HIPAA [29]. Furthermore, the task of maintaining a secure system through Using strong encryption methods and key management protocols is made particularly challenging by insider threats from hostile employees, organizations can meet regulatory mandates for data security, minimize the impact of breaches, and build trust with stakeholder's Common algorithms include AES-256 and RSA for symmetric and asymmetric encryption.

Data Anonymization and Pseudonymization

Pseudonymization and data anonymization are key privacy-preserving techniques used in cloud databases to reduce the risk of identifying individuals while still enabling organizations to process and analyze data. In order to improve data protection and lessen damage in the case of a breach, both are specifically advised by laws like GDPR and HIPAA.

A personal data set is information that might be used to identify an actual individual.

Pseudonymized data is processed personal information that, without the application of extra information, can no longer be linked to a particular bearer [30] .

Data that has been anonymized is personal information that has been handled such that it is difficult to identify or infer details about a particular person.

Access Control and Audit Mechanisms

Another crucial aspect of protecting cloud environments from security risks is access control over data or resources. The permission to access data and resources is granted to authenticated clients, nevertheless, the client should be given instructions on how and what to access. Those permissions are granted by the administrator to the client. It is dependent upon the existing cloud computing techniques, as Role-Based Access Control (RBAC), Attribute-Based Access Control (ABAC), Mandatory Access Control (MAC), and Discretionary Access Control (DAC).

Literature Review

This section presents earlier studies on compliance-driven data governance frameworks in cloud databases, with particular emphasis on GDPR and HIPAA. Table I provides a structured comparison of previous research, focusing on governance strategies, frameworks, and emerging technologies aimed at tackling cloud environments' security, privacy, and regulatory compliance issues.

Singh, Kolluri and Modi (2021) cloud-based database administration platforms. However, this development raises additional security and privacy issues that must be considered. This study looks at the many privacy and security issues with cloud databases, including unauthorized access, data breaches, and non-compliance with rules. Strong encryption tools, data regulation systems, access control tactics, and user sensitivity training. The essay also examines how cloud database administration is affected by compliance regulations like GDPR and HIPAA. The paper outlines procedures that must be adhered to to securely protect personal data within the legal framework; this fact is complemented by the requirement that preventative actions be taken to comply with these recommendations. The purpose of the study is to analyze current issues and investigate potential solutions in order to provide useful information to people and Businesses that rely on database management solutions hosted in the cloud [31] .

et al. (2021) Artificial intelligence (AI) and advanced analytics have revolutionized business processes by increasing efficiency, innovation, and competitiveness. However, risks related to data

security and regulatory compliance have become significant issues as a result of AI-driven procedures being adopted so quickly. These measures have presented remarkable challenges that necessitate effective data governance that includes data protection frameworks to manage these risks and safeguard sensitive information. Data governance and digital transformation, including risk reduction, data governance, and regulatory compliance tactics. AI policy is changing, with a focus on international frameworks like the GDPR. These tactics include data categorization, access control systems, encryption techniques, and real-time audits to improve data security and integrity. Additionally discussed is the need for explainable AI (XAI), which demonstrates how companies may maintain the interpretability of AI models while maintaining regulatory compliance [32].
Shah and Khan (2020) discuss cyberattacks and also make it difficult to enforce stringent privacy regulations because data may occasionally be exchanged with organizations subject to local laws. An individual's privacy may be seriously impacted when private, sensitive information included in an electronic health record is disclosed or made public. Financial losses or social boycotts are two possible outcomes of data leaks. To protect patient data from these threats, a number of privacy rules are in place, including GDPR, HIPAA, and MHR. However, it is becoming increasingly challenging to fully safeguard patient privacy due to the ongoing development of cutting-edge techniques in data analytics, hacking, and ML. A critical analysis of GDPR identified potential areas for improvement, taking into account growing technological use and various secondary uses of EHR [33].

Shastri et al. (2020) Additional rights and safeguards for European people's personal data are offered under the GDPR. Its investigation demonstrates the phenomenon known as the "metadata explosion," which occurs when a significant amount of information must be kept in addition to personal data in order to meet GDPR regulations. Create and implement the GDPR Bench, an open-source benchmark that includes the workloads and metrics required to comprehend and evaluate best practices for database systems that process personal data. It also includes developer recommendations for modifying Redis, PostgreSQL, and a commercial database system to comply with GDPR. These findings have practical ramifications and highlight research difficulties to make GDPR compliance effective in production settings. GDPR compliant systems perform badly on GPDR workloads, and that performance scales poorly as the number of personal data rises [34].

Campbell (2020) As cloud computing grows more and more integrated into contemporary company processes, maintaining security compliance in these settings has become a top priority. The HIPAA and the GDPR set strict guidelines for safeguarding personal and health-related data. Following these guidelines not only keeps companies out of legal hot water but also boosts stakeholder and customer confidence in GDPR and HIPAA solutions. GDPR mandates robust data protection measures, emphasizing data subject rights, breach notification protocols, and strict rules on data processing and transfer-particularly relevant for cloud service providers managing cross-border data flows. HIPAA and customers, and continuous auditing complicate adherence efforts. Adopting strong data governance frameworks and implementing automated compliance monitoring tools, Companies may take advantage of cloud computing's benefits, reduce risks, and safeguard private information [35].

Al-ruithe et al. (2019) Data governance attempts in the past have failed because they were headed by IT and were impacted by inflexible procedures and disjointed actions conducted based on system by system. Governance was mostly informal until recently, and the criteria were broad and ambiguous, working in silos around certain business repositories, and lacking organization-wide support. Even though its significance is widely acknowledged, data governance is still a relatively undeveloped and understudied field. Because data governance research is currently lacking, more research is required to improve practice. The only literature studies that are currently accessible on the topic of data governance are descriptive ones. The present state of data governance research may be understood in a rigorous, logical, and demanding manner thanks to the study's systematic literature review (SLR). In order to help future data governance researchers choose areas in which they may have the most influence, the study aims to offer a reliable conceptual manual [36].

Table 1 Comparative Analysis Of Compliance-Driven Data Governance Literature In Cloud Databases Under Gdpr And Hipaa

Author	Focus Area	Key Findings	Challenges	Future Work
Singh, Kolluri et.al. (2021)	Cloud Database Security & Compliance	Examines privacy/security flaws (e.g., breaches, unauthorized access) in cloud DBMS; recommends encryption, access control, governance, and user training.	Complex compliance with GDPR/HIPAA in multi-tenant environments.	Promote proactive compliance strategies and user awareness.
Onoja et al. (2021)	AI-Driven Digital Transformation & Data Governance	Emphasizes data governance for AI security and GDPR compliance; promotes explainable AI and use of mechanisms	Regulatory uncertainties; need for interpretable AI systems.	Develop XAI tools, improve real-time audit capabilities.
Shah & Khan (2020)	Privacy in Electronic Health Records (EHR)	Highlights privacy risks in EHR sharing; notes GDPR/HIPAA gaps in addressing modern threats like analytics,AI	Regulatory lag behind tech evolution; data misuse risks.	Update and strengthen privacy regulations to reflect AI and secondary data use.

Shastri et al. (2020)	GDPR Compliance in Database Systems	Developed GDPR-bench to assess DBMS compliance; identified metadata explosion and performance issues when handling GDPR workloads.	High performance cost and scalability issues in GDPR-compliant DBMS.	Improve DBMS architectures for scalable GDPR-compliant processing.
Campbell (2020)	Regulatory Compliance in Cloud Computing	Discusses GDPR and HIPAA rules for cloud environments; highlights need for data governance, breach notification, and automated tools for cross-border compliance.	Compliance monitoring, cross-border data issues, and legal complexity.	Automate compliance tracking; strengthen governance frameworks for secure cloud use.
Al-Ruithe et al. (2019)	Examining Data Governance Systematically	Identifies lack of structure and maturity in data governance research and practice; SLR reveals fragmented approaches	Informal governance, fragmented processes, and unclear policies.	Advance structured research on data governance frameworks for enterprise-wide implementation.

Conclusion and Future Work

Cloud databases are cloud-hosted platforms that store and manage sensitive organizational data, enabling accessibility, scalability, and cost efficiency. It is widely adopted in domains such as healthcare and finance, where adherence to cloud database data governance is mandated by regulations like GDPR and HIPAA, with a focus on the legal needs of these regulations. The review identified that while current governance practices such as data classification, retention schedules, anonymization, encryption, and access controls offer significant benefits, they often lack seamless integration and adaptability to Cloud-native systems' dispersed and dynamic nature, including jurisdictional complications and insufficient automation, and emerging privacy risks persist despite the adoption of best practices and standards. Emerging approaches leveraging AI, metadata-driven frameworks, and automated compliance monitoring have demonstrated considerable potential to enhance governance effectiveness, Future studies should concentrate on creating comprehensive, interoperable governance models that include operational scalability and regulatory alignment. Specifically, the design of AI-enabled governance systems capable of adaptive policy enforcement, real-time risk analysis, and predictive compliance monitoring warrants further exploration. Moreover, establishing standardized benchmarks and certification schemes to assess the efficacy of governance frameworks across diverse cloud deployment models is critical. Lastly, creating privacy-preserving technologies like federated learning and differential privacy and promoting cross-jurisdictional harmonization of regulatory requirements can strengthen compliance-driven data governance in cloud environments.

References

[1] S. S, R. R. Patil, and P. C, "Cloud computing an overview," *Int. J. Eng. Technol.*, vol. 7, no. 4, Oct. 2018, doi: 10.14419/ijet.v7i4.10904.

[2] Y. E. Gelogo and S. Lee, "Database Management System as a Cloud Service," *Int. J. Futur. Gener. Commun. Netw.*, vol. 5, no. 2, pp. 71–76, 2012.

[3] S. S. S. Neeli, "The Significance of NoSQL Databases: Strategic Business Approaches and Management Techniques," *J. Adv. Dev. Res.*, vol. 10, no. 1, 2019.

[4] L. A. Tawalbeh and G. Saldamli, "Reconsidering big data security and privacy in cloud and mobile cloud systems," *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 33, no. 7, pp. 810–819, Sep. 2021, doi: 10.1016/j.jksuci.2019.05.007.

[5] H. P. Kapadia, "Voice and Conversational Interfaces in Banking Web Apps," *J. Emerg. Technol. Innov. Res.*, vol. 8, no. 6, pp. g817–g823, 2021.

[6] H. Li, L. Yu, and W. He, "The Impact of GDPR on Global Technology Development," *J. Glob. Inf. Technol. Manag.*, vol. 22, no. 1, pp. 1–6, Jan. 2019, doi: 10.1080/1097198X.2019.1569186.

[7] D. A. Tamburri, "Design principles for the General Data Protection Regulation (GDPR): A formal concept analysis and its evaluation," *Inf. Syst.*, vol. 91, p. 101469, Jul. 2020, doi: 10.1016/j.is.2019.101469.

[8] G. Sartor and F. Lagioia, *The impact of the General Data Protection Regulation (GDPR) on artificial intelligence : study*. 2020.

[9] T. Glenn and S. Monteith, "Privacy in the Digital World: Medical and Health Data Outside of HIPAA Protections," *Curr. Psychiatry Rep.*, vol. 16, no. 11, Nov. 2014, doi: 10.1007/s11920-014-0494-4.

[10] S. S. S. Neeli, "Ensuring Data Quality: A Critical Aspect of Business Intelligence Success," *Int. J. Lead. Res. Publ.*, vol. 2, no. 9, 2021.

[11] H. P. Kapadia, "API-Driven Banking: How COVID-19 Remote Work Boosted Open Banking and Fintech Integrations," *J. Emerg. Technol. Innov. Res.*, vol. 8, no. 10, pp. f514–f519, 2021.

[12] R. N. Zaeem and K. S. Barber, "The Effect of the GDPR on Privacy Policies," *ACM Trans. Manag. Inf. Syst.*, vol. 12, no. 1, pp. 1–20, Mar. 2021, doi: 10.1145/3389685.

[13] V. Singh, "Lessons Learned from Large-Scale Oracle Fusion Cloud Data Migrations," *Int. J. Sci. Res.*, vol. 10, no. 10, pp. 1662–1666, 2021.

[14] A. Balasubramanian, "Building Secure Cybersecurity Infrastructure: Integrating AI and Hardware for Real-Time Threat Analysis," *Int. J. Core Eng. Manag.*, vol. 6, no. 07, pp. 263–271, 2020.

[15] M. Al-Ruithe, E. Benkhelifa, and K. Hameed, "A Conceptual Framework for Designing Data Governance for Cloud Computing," *Procedia Comput. Sci.*, vol. 94, pp. 160–167, 2016, doi: 10.1016/j.procs.2016.08.025.

[16] P. K. Pemmasani and M. Osaka, "Cloud-Based Health Information Systems : Balancing Accessibility with Cybersecurity Risks," vol. 5, no. 2, pp. 22–33, 2019.

[17] A. Thapliyal, P. S. Bhagavathi, T. Arunan, and D. D. Rao, "Realizing Zones Using UPnP," in *2009 6th IEEE Consumer Communications and Networking Conference*, IEEE, Jan. 2009, pp. 1–5. doi: 10.1109/CCNC.2009.4784867.

[18] B. Yuan and J. Li, "The Policy Effect of the General Data Protection Regulation (GDPR) on the Digital Public Health Sector in the European Union: An Empirical Investigation," *Int. J. Environ. Res. Public Health*, vol. 16, no. 6, Mar. 2019, doi: 10.3390/ijerph16061070.

[19] P. H. B. Patel and P. N. Kansara, "Cloud Computing Deployment Models: A Comparative Study," *Int. J. Innov. Res. Comput. Sci. Technol.*, vol. 9, no. 2, pp. 45–50, Mar. 2021, doi: 10.21276/ijircst.2021.9.2.8.

[20] H. J. Bhatti and B. B. Rad, "Databases in Cloud Computing: A Literature Review," *Int. J. Inf. Technol. Comput. Sci.*, vol. 9, no. 4, pp. 9–17, Apr. 2017, doi: 10.5815/ijitcs.2017.04.02.

[21] K. Jones, H. Daniels, S. Heys, A. Lacey, and D. V. Ford, "Toward a Risk-Utility Data Governance Framework for Research Using Genomic and Phenotypic Data in Safe Havens: Multifaceted Review," *J. Med. Internet Res.*, vol. 22, no. 5, May 2020, doi: 10.2196/16346.

[22] S. Rosenbaum, "Data Governance and Stewardship: Designing Data Stewardship Entities and Advancing Data Access," *Health Serv. Res.*, vol. 45, no. 5p2, pp. 1442–1455, Oct. 2010, doi: 10.1111/j.1475-6773.2010.01140.x.

[23] O. R. Grey and R. Brown, "GDPR Compliance: Incident Response and Breach Notification Challenges," in *Cyber Security Practitioner's Guide*, World Scientific, 2020, pp. 275–302. doi: 10.1142/9789811204463_0008.

[24] D. A. Tamburri, "Design principles for the General Data Protection Regulation (GDPR): A formal concept analysis and its evaluation," *Inf. Syst.*, vol. 91, Jul. 2020, doi: 10.1016/j.is.2019.101469.

[25] S. Mbonihankuye, A. Nkunzimana, A. Ndagijimana, and I. García-Magariño, "Healthcare Data Security Technology: HIPAA Compliance," *Wirel. Commun. Mob. Comput.*, 2019, doi: 10.1155/2019/1927495.

[26] W. Moore and S. Frye, "Review of HIPAA, Part 1: History, protected health information, and privacy and security rules," *J. Nucl. Med. Technol.*, vol. 47, no. 4, pp. 269–272, 2019, doi: 10.2967/JNMT.119.227819.

[27] I. Keshta and A. Odeh, "Security and privacy of electronic health records: Concerns and challenges," *Egypt. Informatics J.*, vol. 22, no. 2, pp. 177–183, Jul. 2021, doi: 10.1016/j.eij.2020.07.003.

[28] M. A. Adoyo, "Landscape analysis of healthcare policy: the instrumental role of governance in HIV/AIDS services integration framework," *Pan Afr. Med. J.*, vol. 36, May 2020, doi: 10.11604/pamj.2020.36.27.22795.

[29] W.-B. Lee and C.-D. Lee, "A Cryptographic Key Management Solution for HIPAA Privacy/Security Regulations," *IEEE Trans. Inf. Technol. Biomed.*, vol. 12, pp. 34–41, 2008, doi: 10.1109/TITB.2007.906101.

[30] J. F. Marques and J. Bernardino, "Analysis of data anonymization techniques," *IC3K 2020 - Proc. 12th Int. Jt. Conf. Knowl. Discov. Knowl. Eng. Knowl. Manag.*, vol. 2, no. Ic3k, pp. 235–241, 2020, doi: 10.5220/0010142302350241.

[31] A. Singh, S. Kolluri, and T. B. Modi, "Security and Privacy Challenges in Cloud-Based Database Management : Strategies and Solutions," *J. Technol. Manag.*, vol. 01, no. 01, pp. 32–40, 2021.

[32] J. P. Onoja, O. Hamza, A. Collins, U. B. Chibunna, A. Eweja, and A. I. Daraojimba, "Digital Transformation and Data Governance: Strategies for Regulatory Compliance and Secure AI-Driven Business Operations," *J. Front. Multidiscip. Res.*, vol. 2, no. 1, pp. 43–55, 2021, doi: 10.54660/.ijfmr.2021.2.1.43-55.

[33] S. M. Shah and R. A. Khan, "Secondary use of electronic health record: Opportunities and challenges," *IEEE Access*, vol. 8, pp. 136947–136965, 2020, doi: 10.1109/ACCESS.2020.3011099.

[34] S. Shastri, V. Banakar, M. Wasserman, A. Kumar, and V. Chidambaram, "Understanding and benchmarking the impact of GDPR on database," *Proc. VLDB Endow.*, vol. 13, no. 7, pp. 1064–1077, 2020, doi: 10.14778/3384345.3384354.

[35] L. Campbell, "Security Compliance in Cloud Computing (e.g., GDPR, HIPAA)," 2020.

[36] M. Al-ruithe, S. Arabia, E. Benkhelifa, and K. Hameed, *A systematic literature review of data governance and cloud data governance A Systematic Literature Review of Data Governance & Cloud Data Governance*, no. December. 2019. doi: 10.1007/s00779-017-1104-3.