# Data Governance and Security

**A Project Work Synopsis**

*Submitted in the partial fulfilment for the award of the degree of*

**BACHELOR OF ENGINEERING**

**IN**

**COMPUTER SCIENCE WITH SPECIALIZATION IN**

**BIG DATA AND ANALYTICS**

**Submitted by:**

20BCS3908 Kanwar Arinjai Singh  
20BCS3873 Abhinav Sharma  
20BCS3945 Anupam Kumar

Under the Supervision of:

**Navjeet Kaur (E16069)**



**CHANDIGARH UNIVERSITY, GHARUAN, MOHALI - 140413,**

**PUNJAB**

**September 2022**

# Abstract

Data governance is the set of policies, processes, and procedures that define how data is managed within an organization. It ensures that data is used and shared in a consistent and compliant manner, and that it is protected from unauthorized access, use, or disclosure. Data security is the protection of data from unauthorized access, use, disclosure, disruption, modification, or destruction. It is a critical part of data governance, and it helps to ensure the confidentiality, integrity, and availability of data. Data governance and security are closely related, and they are both essential for protecting data. A comprehensive data governance framework should include security measures to protect data, and a strong data security program should be aligned with the organization's data governance policies and procedures. Data governance and security framework, organizations can help to protect their data from unauthorized access, use, or disclosure. This can help to prevent data breaches, protect the privacy of individuals, and comply with applicable laws and regulations.

# Table of Contents

|  |  |
| --- | --- |
| Title Page | i |
| Abstract | ii |
| 1.    Introduction | 1 |
| 1.1  Problem Definition |  |
| 1.2 Project Overview |  |
| 1.3 Hardware Specification |  |
| 1.4 Software Specification |  |
| 2.    Literature Survey | 3 |
| 2.1 Existing System |  |
| 2.2 Proposed System |  |
| 2.3 Literature Review Summary |  |
| 3.    Problem Formulation | 5 |
| 4.    Research Objective | 6 |
| 5.    Methodology | 8 |
| 6.    Conclusion | 10 |
|  |  |
|  |  |
|  |  |

# 1. INTRODUCTION

## 1.1 Problem Definition

The project's primary focus is on addressing challenges related to data governance and security within the organization's data ecosystem. The proliferation of data sources, increasing data volumes, and the need to ensure data privacy and compliance have led to several critical issues that require resolution. The organization lacks a robust system for maintaining data quality and consistency across various data sources and databases. Inaccurate, incomplete, or inconsistent data hampers decision-making processes and undermines trust in the data. There is a lack of proper access controls and permissions management for data stored within the organization. Unauthorized access to sensitive data poses a significant security risk and could lead to breaches of confidentiality. The organization struggles to meet data privacy regulations and industry-specific compliance standards (such as GDPR, HIPAA, etc.). Non-compliance could result in legal penalties and damage to the organization's reputation.

**1.2 Problem Overview**

In today's data-driven landscape, organizations face significant challenges in ensuring the effective management, security, and compliance of their data assets. These challenges stem from the growing volume and complexity of data, evolving privacy regulations, and the need to maintain data accuracy and integrity. The project seeks to enhance data accuracy and consistency by establishing standardized data entry processes, validation mechanisms, and data quality metrics. This ensures that decision-makers rely on trustworthy and consistent data. By implementing robust access control mechanisms, the project aims to prevent unauthorized access to sensitive data. Role-based access and encryption techniques will be employed to safeguard information from breaches and unauthorized use. The project addresses the challenge of complying with data protection regulations such as GDPR, HIPAA, and others. This involves creating processes and policies that align with these regulations to avoid legal repercussions and maintain customer trust.

**1.3 Hardware Specification**

1. RAM: 4Gb minimum
2. Intel i5 processor or equivalent
3. 256gb SSD storage

# 2. LITERATURE SURVEY

## 2.1 Existing System

Master Data Management (MDM) tools are commonly used in data governance projects, to define a business glossary which is a single point of reference for critical business data.

## 2.2 Proposed System

Data Governance and Security aims to comprehensively address the limitations and challenges of the existing system. It introduces a cohesive framework that encompasses policies, processes, technologies, and cultural changes to ensure effective data governance, enhanced security, and regulatory compliance.

# 2.3 Literature Review Summary

Data governance and security are critical components of modern organizations, ensuring the effective management, protection, and responsible use of data. Scholars emphasize that data governance is essential for maintaining data quality, consistency, and integrity. Organizations need to establish clear roles, responsibilities, and processes for data stewardship to ensure accurate and reliable data. The literature highlights the increasing challenges of data breaches, cyber-attacks, and unauthorized access. Data security measures must evolve to address new threats, such as ransom ware and advanced persistent threats Research underscores the significance of regulatory compliance, including GDPR, HIPAA, and others. Non-compliance can lead to substantial fines and reputational damage. Organizations must align data governance and security practices with these regulations. Scholars emphasize the adoption of data classification frameworks to categorize data based on sensitivity. This approach enables tailored security measures and appropriate access controls for different data types.

# 3. PROBLEM FORMULATION

The Data Governance and Security project aims to address the challenges and deficiencies within the organization's data management practices that impede effective data governance and compromise data security. In today's data-driven landscape, organizations rely on data to make critical decisions and drive innovation. However, the proliferation of data sources, evolving data protection regulations, and the increasing sophistication of cyber threats have highlighted significant gaps in the organization's data governance and security framework. The organization currently lacks a systematic and cohesive approach to data management. Data is stored across disparate systems, leading to data silos, inconsistencies, and a lack of centralized control. The existing data security measures are fragmented and do not provide comprehensive protection against modern cyber threats. Unauthorized access, data breaches, and cyber-attacks remain significant risks due to the absence of robust access controls and monitoring mechanisms.

# 4. OBJECTIVES

Data Governance and Security project are designed to address the challenges identified in the existing system and to establish a robust framework for data governance, security, and compliance.

1. **Enhance Data Quality and Consistency:** Implement processes and mechanisms to ensure data accuracy, consistency, and integrity across all data sources and systems.
2. **Establish Strong Access Controls:** Develop and deploy robust access control measures, including role-based access and authentication mechanisms, to prevent unauthorized data access.
3. **Ensure Regulatory Compliance:** Align data practices with relevant data protection regulations (e.g., GDPR, HIPAA) and industry standards to avoid legal penalties and maintain compliance.
4. **Define Data Ownership and Accountability:** Clearly define roles and responsibilities for data ownership and stewardship to ensure accurate data management and accountability.
5. **Optimize Data Lifecycle Management:** Establish well-defined data lifecycle management processes, including retention, archiving, and disposal, to minimize storage costs and adhere to regulatory requirements.
6. **Implement Data Classification Framework:** Introduce a data classification framework that categorizes data based on its sensitivity, enabling tailored security measures for different types of data.
7. **Enable Real-Time Monitoring and Auditing:** Implement continuous data monitoring and auditing mechanisms to promptly detect and address unauthorized access and unusual activities.

# METHODOLOGY

Data Governance and Security project involves a structured approach to implementing the proposed solutions and achieving the defined objectives. The methodology encompasses several phases, each with specific activities and tasks.

**1. Project Initiation:**

Define project scope, objectives, and stakeholders.

Establish a project team with representatives from IT, data management, legal, compliance, and other relevant departments.

Secure executive sponsorship and allocate necessary resources.

**2. Current State Assessment:**

Conduct a thorough assessment of the existing data governance and security practices.

Identify weaknesses, vulnerabilities, and compliance gaps in the current system.

Gather data on data flows, storage systems, access controls, and compliance efforts.

**3. Requirements Gathering:**

Collaborate with stakeholders to gather specific requirements for data governance, security, and compliance.

Define the scope of the data classification framework, access control mechanisms, compliance standards, and monitoring needs.

1. **Framework Design:**

Design a comprehensive data governance framework that includes policies, procedures, and roles.

Develop a data classification framework, defining data categories and corresponding security measures.

Create a role-based access control model that aligns with the organization's structure and data needs.

Formulate data retention, archiving, and disposal policies.

**5. Technology Implementation:**

Select and implement appropriate technological solutions to support the data governance and security framework.

Integrate access control mechanisms, encryption, monitoring tools, and auditing systems.

Develop or configure data loss prevention (DLP) systems to prevent data leakage.

# 6. CONCLUSION AND FUTURE SCOPE

**Future Scope**

**Emerging Technologies:** As new technologies emerge, such as quantum computing and AI-driven attacks, the organization must stay vigilant and adapt security measures to address these evolving threats.

**Continuous Monitoring and Improvement:** Implement a continuous monitoring strategy to detect and respond to emerging security threats and vulnerabilities in real-time. Regularly update and refine security measures and policies to stay ahead of potential risks.

**Advanced Data Analytics:** Utilize advanced analytics to identify patterns and anomalies in data access and usage. This can enhance the organization's ability to detect unusual activities that may indicate security breaches.

**Incident Response Plan:** Develop a robust incident response plan to ensure a coordinated and effective response in the event of a data breach or security incident. Regularly conduct simulated exercises to test the plan's effectiveness.

**Blockchain and Data Integrity:** Explore the integration of blockchain technology to ensure data integrity and tamper-proof auditing of data access and changes.

**Global Data Privacy Regulations:** Stay informed about emerging global data privacy regulations and adapt the framework to ensure compliance across different jurisdictions.

**Vendor and Third-Party Management:** Strengthen the assessment and monitoring of third-party partners' data security practices to prevent potential breaches through external collaborations.

**Ethical Data Use:** Extend the framework to include ethical considerations in data usage, ensuring that data is utilized in ways that align with the organization's ethical principles.

**Automated Compliance:** Explore automation solutions that help streamline compliance efforts, making it easier to adapt to changing regulations.

**Employee Empowerment:** Continue to invest in employee training and awareness programs to ensure a sustained culture of data responsibility and security consciousness.

**Conclusion:**

In conclusion, the Data Governance and Security project has successfully addressed the existing challenges within the organization's data ecosystem. By implementing a comprehensive framework encompassing policies, processes, technologies, and cultural changes, the project has significantly improved data governance, security, and compliance practices. The project's outcomes have led to enhanced data accuracy, minimized security vulnerabilities, and a heightened awareness of responsible data management practices among employees. The project's success is evidenced by the establishment of a robust data classification framework, the implementation of strong access controls, the integration of security measures across the data lifecycle, and the adoption of a proactive approach to regulatory compliance. Employee data literacy has improved, leading to a workforce that actively contributes to data protection efforts. Additionally, the project has fostered a cultural shift toward data responsibility, embedding security awareness and compliance into the organization’s values.

**Conclusion:**

1. Abraham, R., Schneider, J., & vom Brocke, J. (2019). Data governance: A conceptual framework, structured review, and research agenda. International Journal of Information Management, 49, 424-438
2. Al-Badi, A., Tarhini, A., & Khan, A. I. (2018). Exploring big data governance frameworks. In Procedia Computer Science (Vol. 141, pp. 271–277). Elsevier B.V. <https://doi.org/10.1016/j.procs.2018.10.181>
3. Al-Ruithe, M., Benkhelifa, E., & Hameed, K. (2019). A systematic literature review of data governance and cloud data governance. Personal and Ubiquitous Computing, 23(5-6), 839-859.
4. Cheong, L. K., & Chang, V. (2007). The need for data governance: a case study. ACIS 2007 Proceedings, 100.
5. Cichy, C., & Rass, S. (2019). An overview of data quality frameworks. IEEE Access, 7, 24634-24648.