IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Enhancing Compliance And Security In Cloud-Based Data Engineering For Financial Data Lakes Through Automated Data Governance

Venkatakota Sivakumar Kopparapu
NTT DATA Americas, Inc

Abstract

Automated data governance's role in helping cloud based financial data lakes achieve compliance and security is researched. The research addresses the way automated systems handle challenges such as data quality, privacy and regulatory compliance. The scalability of disposal systems is also authorized to help such systems adapt to more and more volumes of data and exchange in regulations. There is less human error involved in data disposal and it continues to be monitored mechanically and approach is inside limits in terms of ensuring data security. The research determines recommended practices for implementing automated disposal frameworks that prioritized maximal data shelter and regulative compliance. Maintaining the unity and credentials of the fiscal data are authorized in the ever-changing regulative environs and these are authorized practices.

Keywords: financial data lakes, automated data governance, data quality, compliance, security, scalability, governance frameworks, privacy, data protection, regulatory compliance.

INTRODUCTION

Cloud based data engineering has covered the way in changing the work of fiscal data management. Financial organizations are progressively using data lakes to store large volumes of organized and unstructured data. Challenges such as the power to check entry with strict regulations and data credentials in these environments, are significant. The result of this job is with automated data disposal that streamlines data monitoring, check over approach and entry with applicative regulations. It investigates the way automated disposal mechanisms can improve entry and credentials in cloud fiscal data lakes.

Aim

The aim of this project is to look at the way automated data governance improves compliance and security in cloud-based data engineering for financial data lakes.

Objectives

- To identify the primary difficulties of guaranteeing compliance and security in cloud-based financial data lakes
- To examine the role of automated data governance in simplifying regulatory compliance and security procedures
- To investigate the effects of automated governance on data quality and privacy in financial data management
- To recommend best practices for automating data governance to improve security and compliance in data lakes

Research Questions

- What are the main problems guaranteeing compliance and safety in cloud-based financial data lakes?
- How can automatic data governance help to safeguard data quality and privacy in financial schemes?
- What role does automation play in satisfying regulatory requirements for cloud-based data manufacturing in finance?
- What are the best approaches for adopting automatic data governance to improve safety and compliance?

RESEARCH RATIONALE

Compliance and security concerns about cloud-based data engineering in financial institutions have been

increased by its rapid adoption. Financial data lakes are required to adhere strictly to the regulatory guidelines and bolster data protection measures to protect sensitive information stored in them. However, data management across distributed systems can be complex and lead to security gaps and noncompliance [1]. Manual governance techniques are no longer sustainable as data quantities and regulatory expectations grow. Timely solution of automated data governance provides ongoing monitoring, access control and compliance with the regulation.

LITERATURE REVIEW

Challenges in Ensuring Compliance and Security in Cloud-Based Financial Data Lakes

Deploying financial data lakes on cloud platforms has many challenges especially in terms of compliance and security. Large amounts of sensitive data to be managed have to comply with strict regulatory requirements which have to be managed by the financial institutions. Cloud environments are very complex and the data architecture as well, there are often challenges keeping data consistent as to data security or compliance. Cloud basis is fragmented and hard to check the approach for any data liaison that increases the initiative of unauthorized data approach [2]. Cloud environs is energizing and continues to develop from the manufacture perspective, necessitating the need of successive monitoring and fitting towards complying with manufacture appropriate regulations.

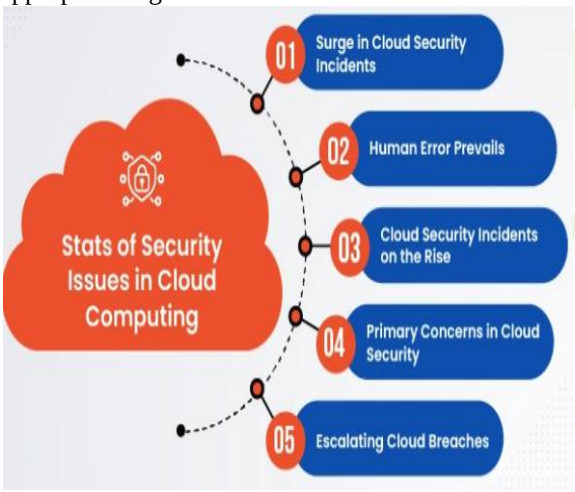

Fig 1: Security issues in Cloud Computing

Integrating many data sources and formats into fiscal data lakes can result in data discrepancies, peradventure leading to entry difficulties. Handling data contained personally recognizable data PII or fiscal transactional data made sharing exceedingly dirty and maintaining data privateness passim this ferment can be important for entry and credentials [3]. The absence of standardized disposal mechanisms in the cloud hinders administrator attempts. These challenges fit dirty to hold a consistent as well as fix and manageable data direction example for the fiscal institutions.

The Role of Automated Data Governance in Regulatory Compliance and Security Procedures

Cloud based automated data disposal has meaningful grandness in regulative entry and ensures the credentials processes in the cloud based fiscal data lakes. Automations of disposal tasks in fiscal institutions enable them to declare human error and maintain consistent consistency in the data superintendence [4]. Continuous approach check was implemented and only authorized workers have access to important data owing to automated systems that perpetually tracked staff access. The same systems make it easy for institutions to operate in real-time on evolving regulatory standards, like GDPR or Dodd-Frank.

Automated governance technologies can automatically update and implement security rules to reduce the likelihood of regulatory infractions. This makes sure that all data activities are transparent and the financial institutions have a transparent audit trail of all that went on. Those are important for regulatory reporting and internal audits. Proactive data security is another way that automated data governance also supports proactive data security by identifying and addressing vulnerabilities before any of them become breaches [5]. The technologies apply data protection standards across numerous cloud environments while maintaining a uniform security strategy.

Impact of Automated Data Governance on Data Quality and Privacy in Financial Data Management

Data quality and privacy within financial data management get greatly affected with automated data governance. It helps to guarantee data consistency, ensuring that the data is stored in the financial data lake, the data is also being monitored and validated all the time. Real time data inconsistencies and errors can be detected and rectified by automated systems leading to improvement in data quality across the terms [6]. The financial institutions can use accurate and high-quality data that can assist them in making better informed decisions as seen above and reduce the risk of compliance violation because of data discrepancy. Automated data governance limits the access of sensitive financial information to only the users who are duly authorized in terms of privacy.

Fig 2: Automated Data Governance

Automated governance systems are good at tracking and monitoring data access activities through an audit trail to guarantee accountability. Data governance in an automated fashion is able to provide a solid mechanism to manage data quality and privacy protection that are important at the highest risk levels in the highly regulated financial sector [7]. The automation of governance lowers the likelihood of any

human error in the time it comes to handling sensitive information and also maximizes compliance with privacy standards and security in general.

Best Practices for Implementing Automated Data Governance to Enhance Security and Compliance in Data Lakes

Improving the security and compliance of data lakes requires data governance to be implemented in a strategic way that can be used for automating. Clear disposal policies can be set up that are aligned to the manufacture regulations guiding fiscal institutions. Policies can be enforced mechanically as well on automated systems, allowing entry with legal and regulatory requirements to proceed successive [8]. Access check is a very authorized broker to be implemented.

User approach rights can be habitually automated verified so that only those having authorized approach to live data can interact with it. Data masking and encryption technologies are combined to ensure that sensitive data remains private and remains inaccessible to persons not authorized to view them. Automated tools of auditing and monitoring of the data activities can be adopted by institutions [9]. For example, automated governance systems also help in identifying potential gaps of compliance through regular audits. Automated systems can scale to meet growing data quantities and changing regulatory requirements. This adaptability guarantees that the data governance architecture stays strong as the data landscape and compliance needs evolve [10]. Financial institutions can optimize data security, privacy and compliance of data in data lakes with these best practices.

Literature gap

There is a considerable knowledge gap about the complete impact of automated data governance on data quality in financial data lakes. Proofs of scalability of automated governance systems across different cloud platforms are not available. These deficiencies impede the creation of comprehensive frameworks for improving security, compliance and data quality in large-scale financial systems. The research of manual data governance and general cloud data management dominates accepted papers while little research exists about automated data governance specifically applied to financial data lakes. The financial sector lacks sufficient research which analyzes the effects that automated systems have on regulatory compliance requirements and data security as well as privacy protection measures.

METHODOLOGY

The research depends on qualitative methods that use **secondary sources** to examine automated data governance methods that improve compliance and security in cloud-based financial data lakes. The research aims to examine complex human conduct and financial sector implementations of organizational practices and technology through adopting an **interpretivist philosophy**. The analyzed method suits the investigation of automated governance and compliance relationships operating in cloud-based environments.

The research employs **deductive approach** that use previous theories of data governance and regulatory compliance and cloud data management to evaluate secondary data. The research starts by developing its conceptual framework through literature analysis then tests these concepts by studying financial institutions that adopted automated data governance in their operations [11]. Through this research the scholar can establish the validity of previous theoretical frameworks while building valuable knowledge for modern research paradigms.

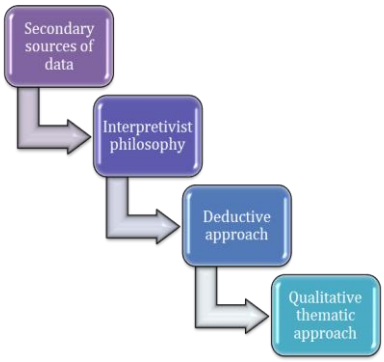

Fig 3: Methodology

They conduct a **thematic analysis** on the collected data to detect repeated patterns which demonstrate the way automated governance affects data protection along with confidentiality rules and conformity guidelines. The study can break down its data into thematic categories which include elements regarding compliance difficulties together with security matters and governance excellence. The researcher uses secondary data obtained from academic journals along with industry reports and white papers and case studies from reputable sources [12]. By employing this research method both validity and comprehensive insight into automated data governance for financial data lake security and compliance can be achieved.

DATA ANALYSIS

Theme 1: Automated data governance enhances compliance by streamlining regulatory reporting and reducing manual errors.

Automated data governance systems work as essential compliance tools since they deliver automated capabilities to handle regulatory requirements. The multitude of regulations which financial institutions operate under includes **GDPR** and **SOX** and **Basel III** because these demand accurate report delivery [13]. The practice of manual reporting introduces avoidable human mistakes while operating slowly which sometimes results in regulatory penalties. Financial institutions enhance the speed of their compliance processes by utilizing automated governance tools that operate on exact time-sensitive data collection and conduct both validation operations and reporting functions. The implementation of automation eliminates human mistakes during data entry and avoids late reporting that enables institutions to comply with relevant legal requirements [14]. The tools provide institutional support through automatic

adaptation of new regulatory rules which enables them to incorporate these changes into their governance structures. Financial institutions demonstrate their dedication to accurate reporting operations through secure processes which results in better regulatory compliance and reduced costs of fines and improved relationships with regulatory bodies [15]. This also helps in faster adaptation to change in the regulatory environment, whereby compliance frameworks of financial institutions can be updated in real time using automated tools.

Theme 2: Data security is fortified by automated governance tools that enforce real-time access controls and encryption protocols.

The protection of financial data lakes happens thanks to automated systems which both determine access controls and maintain encryption throughout this system constantly. Cloud-based environments which contain massive amounts of sensitive financial data require absolute priority for establishing solid security measures. Automated governance tools enable real-time access monitoring which confirms that exclusive authorized staff members obtain access to critical information assets. Data lake security increases substantially through **RBAC** implementation as well as encryption standard enforcement of entire data lake [16]. The system applies encryption protocols through automation without any regard to data location or status when data is moving or stored. The human effort required for manual security maintenance becomes too complex because of the enormous amount and distinctive nature of data present. Automated governance solutions in financial institutions use standardized protective measures to reduce security risks generated by human operations [17]. It speeds up response to security threats so that potential damage can be reduced.

Theme 3: Automated governance improves data privacy by managing consent and ensuring compliance with privacy regulations.

Financial organizations strongly prioritize data privacy because GDPR and other privacy rules keep becoming more complicated. The management of consent together with preferences under automated data governance systems protects privacy by respecting all privacy rights. Financial institutions need to exercise proper control over personal data management as data collection grows because users require both consent and documented approval. Tools run by automation enable firms to track and handle consent usage real-time for individual data points so authorized activities maintain compliance [18]. These systems offer clear visibility because users can view and modify their privacy settings without any difficulty. Institutions protect themselves from privacy law non-compliance risks when they automate consent management because this ensures regulatory requirements persist. The implementation of automated governance systems allows institutions to track and implement data retention policies thus preventing unnecessary data storage. **Financial institutions** achieve operational efficiency and data privacy protection at a high level through process automation for customers that builds their trust and reduces privacy risks. Additionally, these systems also enable automation of revocations of consent, so that opting out of the use of an individual's data is always processed, either deleting or anonymizing the data while necessary [19]. This further bolster privacy measures, making the same rule apply to users' rightful rights.

Theme 4: Automated data governance facilitates continuous monitoring and auditing, ensuring consistent compliance with regulatory requirements.

The main benefit of automated data governance systems enables financial data monitoring at all times accompanied by auditing capabilities. Compliance requirements have mandated organizations to perform ongoing oversight where they maintain thorough documentation about usage logs and regulatory changes for monitoring purposes. The automatic audit trail system tracks all data events to make compliance audits more transparent because such systems operate independently for this purpose [20]. **Automated governance systems** enable financial institutions to achieve faster and more effective result production which reduces their manual audit costs. Continuous monitoring of the system lets organizations detect regulatory breaches immediately to take appropriate corrective measures quickly. Organizations obtain automation benefits in governance tools because they use pre-programmed procedures that align organization policies with changing regulations. Continuous regulatory compliance under automated data governance cuts down excessive work for compliance teams and results in improved operational efficiency as well as sustained compliance to prevent penalties and reputation damage [21]. The financial institutions become audit ready all time and reduce loading to compliance team by automating compliance audits.

FUTURE DIRECTIONS

Newsletter authors ought to research the way integrating artificial intelligence (AI) and blockchain technology with automatic data governance systems can enhance security standards in cloud-based financial data lakes for future. **Blockchain technology** provides transparent data management systems that also maintain unalterable records for enhancing data integrity [22]. AI enhances predictive abilities of governance tools to spot risks that can occur ahead of time.

CONCLUSION

Conclusively, there is a crucial necessity to get rid of data errors and security failures in policy compliance by using automated data management in cloud financial data lakes. Concerns of timely and accurate regulatory reporting, implementation and monitoring of controls, and information security are some of fundamental issues that financial institution have to

deal with. These are following benefits of automation: automation minimizes risk of human errors, increases quality of data, and provides compliance with current and new standards. The role of automated governance remains important because of this and since more financial institutions are adopting cloud technology.

References

[1] Asante, M., Epiphaniou, G., Maple, C., Al-Khateeb, H., Bottarelli, M. and Ghafoor, K.Z., 2021. Distributed ledger technologies in supply chain security management: A comprehensive survey. *IEEE Transactions on Engineering Management*, *70*(2), pp.713-739.

[2] Katari, A. and Ankam, M., 2022. Data Governance in Multi-Cloud Environments for Financial Services: Challenges and Solutions. *Educational Research (IJMCER)*, *4*(1), pp.339-353.

[3] Olabanji, S.O., Oladoyinbo, O.B., Asonze, C.U., Oladoyinbo, T.O., Ajayi, S.A. and Olaniyi, O.O., 2024. Effect of adopting AI to explore big data on personally identifiable information (PII) for financial and economic data transformation. *Available at SSRN 4739227*.

[4] Alao, O.B., Dudu, O.F., Alonge, E.O. and Eze, C.E., 2024. Automation in financial reporting: A conceptual framework for efficiency and accuracy in US corporations. *Global Journal of Advanced Research and Reviews*, *2*(02), pp.040-050.

[5] Nookala, G., 2024. Adaptive Data Governance Frameworks for Data-Driven Digital Transformations. *Journal of Computational Innovation*, *4*(1).

[6] Ghelani, H., 2024. AI-Driven Quality Control in PCB Manufacturing: Enhancing Production Efficiency and Precision. *Valley International Journal Digital Library*, pp.1549-1564.

[7] Nookala, G., 2024. Adaptive Data Governance Frameworks for Data-Driven Digital Transformations. *Journal of Computational Innovation*, *4*(1).

[8] Parycek, P., Schmid, V. and Novak, A.S., 2024. Artificial Intelligence (AI) and automation in administrative procedures: Potentials, limitations, and framework conditions. *Journal of the Knowledge Economy*, *15*(2), pp.8390-8415.

[9] Elumilade, O.O., Ogundeji, I.A., Ozoemenam, G.O.D.W.I.N., Achumie, H. and OMOWOLE, B.M., 2024. Advancing audit efficiency through statistical sampling and compliance best practices in financial reporting. *IRE Journals*, *7*(9), pp.434-437.

[10] Rachakatla, S.K., Ravichandran, P. and Kumar, N., 2022. Scalable Machine Learning Workflows in Data Warehousing: Automating Model Training and Deployment with AI. *Australian Journal of AI and Data Science*.

[11] Machado Ribeiro, V.H., Barata, J. and da Cunha, P.R., 2022. Sustainable data governance: a systematic review and a conceptual framework.

[12] Jensen, E.A., Wong, P. and Reed, M.S., 2022. How research data deliver non-academic impacts: A secondary analysis of UK Research Excellence Framework impact case studies. *PloS one*, *17*(3), p.e0264914.

[13] AttaYeboah, N.O., 2022. A comparative analysis of South African and Chinese System of Corporate Governance: A case of Standard Bank South Africa and Agricultural Bank of China.

[14] Matarneh, S., Elghaish, F., Rahimian, F.P., Dawood, N. and Edwards, D., 2022. Automated and interconnected facility management system: An open IFC cloud-based BIM solution. *Automation in Construction*, *143*, p.104569.

[15] Miglionico, A., 2022. The use of technology in corporate management and reporting of climate-related risks. *European Business Organization Law Review*, *23*(1), pp.125-141.

[16] Kothandapani, H.P., 2022. Optimizing Financial Data Governance for Improved Risk Management and Regulatory Reporting in Data Lakes. *International Journal of Applied Machine Learning and Computational Intelligence*, *12*(4), pp.41-63.

[17] Katari, A. and Ankam, M., 2022. Data Governance in Multi-Cloud Environments for Financial Services: Challenges and Solutions. *Educational Research (IJMCER)*, *4*(1), pp.339-353.

[18] van Groesen, W. and Pauwels, P., 2022. Tracking prefabricated assets and compliance using quick response (QR) codes, blockchain and smart contract technology. *Automation in Construction*, *141*, p.104420.

[19] Jesus, V. and Pandit, H.J., 2022. Consent receipts for a usable and auditable web of personal data. *IEEE Access*, *10*, pp.28545-28563.

[20] Tysebaert, S., Lloreda Sanchez, H., Riviere, E. and Rath, A., 2022. AuditTrust: A blockchain-based secure audit trail for data sharing in distributed environment. *Commun. Comput. Inf. Sci*.

[21] Sharma, C. and Vaid, A., 2022. Converging SAP, AI, and data analytic for transformative business management. *World J. Adv. Res. Rev*, *14*(3), pp.736-761.

[22] Shaik, M. and Bojja, G.R., 2022. *Advanced Identity Access Management and Blockchain Integration: Techniques, Protocols, and Real-World Applications for Enhancing Security, Privacy, and Scalability in Modern Digital Infrastructures*. Libertatem Media Private Limited.