AMERICAN JOURNALS
Open Access, Peer Reviewed Journals

American Journal of Technology and Applied Sciences
ISSN (E): 2832-1766 Volume 38, July - 2025

PROBLEMS AND PROSPECTS OF INTEGRATION OF DATA GOVERNANCE WITH MULTICLOUD PLATFORMS

Alisher Serikbayev,
IT Expert, USA

ABSTRACT

This article examines the key issues and prospects for integrating data management practices (Data Governance) in multi-cloud environments. The main challenges arising in the context of multi-cloud architectures are identified and systematized, including technical, organizational and legal aspects. An analysis of the impact of multi-clouds on data management processes is conducted, and approaches to overcoming the identified difficulties are proposed. In conclusion, strategic directions for further research and practical development in this area are outlined, which allows for the formation of a comprehensive vision of solving the problem of ensuring data integrity and security in distributed cloud infrastructures.

KEYWORDS

Data Governance, Multi-Cloud, multicloud platforms, Policy-as-Code (PaC), Open Policy Agent (OPA), Zero Trust, Data Mesh, Federated Governance, Cloud Repatriation, Gaia-X, AI-Driven Governance, Edge Computing, FinOps, Secure Multi-Party Computation (MPC), Fully Homomorphic Encryption (FHE), compliance in cloudy environments, sovereignty data.

Introduction

In the enterprise space, there is growing interest in multi-cloud strategies, which involve using services from multiple cloud service providers (public, private, hybrid). According to ISO/IEC 22123-1, multi-cloud is defined as a deployment model that uses more than one cloud service from different providers [1].

The use of multi-cloud environments is due to a number of strategic advantages:

1Architectural stability. Failure of one provider does not lead to complete system inoperability, which ensures high fault tolerance and continuity of business processes.

2. Flexibility and specialization. Multi-cloud environments allow you to use each provider's unique, specialized services, for example, for processing big data or machine learning.

3. Avoiding dependence on a single vendor lock-in). Distributing the load and data across multiple providers reduces the risk of becoming dependent on a single vendor , which is especially important in the case of long-term contracts and the use of proprietary APIs [2].

Despite the obvious benefits, multi-cloud environments pose serious challenges for data management (Data Governance):

1. Fragmentation and complexity of management. Each provider offers its own tools, security models and metadata formats, which complicates control, audit and creates «dark areas» in access and data management.
2. Security issues. In the scientific article «Cybersecurity Risks in Multi-Cloud Environments» identified increased risks: mismatched authentication mechanisms, different levels of protection, API vulnerabilities and new attack vectors associated with a complex structure [3].
3. Regulatory compliance. The distribution of data across multiple jurisdictions makes it difficult to track and requires careful consideration of its storage, encryption, and cross-border transfer, especially given the requirements of GDPR, HIPAA, and other local regulations.
4. Policy consistency. The diversity of providers leads to differences in access, classification, encryption, and SLA policies. In the «Security issues in cloud environments: a survey» emphasizes the need for a hybrid security model where compliance policies are integrated into a multi-cloud architecture [4].

According to forecasts by the American research company «Gartner», multi-cloud strategies will be actively implemented in more than 80% of large organizations. However, as noted by a group of scientists who published the article «A Survey Study on Major Technical Barriers Affecting the Decision to Adopt Cloud Services», security, privacy and data portability concerns are the main barriers to moving to the cloud [5]. Therefore, achieving a balance between the benefits of multicloud and the risks is only possible with a competent approach to Data Governance , which integrates political, technical and regulatory aspects into a single strategy adapted to a heterogeneous cloud environment.

Integration of data management practices (Data Governance) in multi-cloud environments presents a number of fundamental challenges that make it difficult to build a holistic, secure, and compliant governance system.

First, platform fragmentation and policy incompatibility. The diversity of cloud providers leads to the creation of a heterogeneous ecosystem, where each provider uses its own management tools, access models, metadata formats, and APIs. This significantly reduces interoperability and creates the risk of conflicting security policies. According to experts, interoperability and data portability are key challenges that are especially aggravated in multi-cloud scenarios due to the presence of conflicting security policies [6].

Secondly, data fragmentation and shadow IT. Data fragmentation occurs due to the use of disparate storage and uncontrolled access channels, known as shadow IT. This complicates monitoring and control, which leads to a loss of data visibility and increased security risks. In addition, as noted by Folio3 Software, the uncontrolled dissemination of data outside of corporate IT control increases the likelihood of leaks and reduces the reliability of information [7].

Third, data visibility and lineage. The lack of a single catalog or centralized system for tracking the lineage of data lineage) between different platforms seriously hampers quality management, auditing and compliance. Privately held Maruti Techlabs, which provides a wide range of IT services and technology solutions, highlights the challenges associated with scaling, dynamic changes and the lack of unified tools for monitoring data flows in a multi-cloud environment [8].

Fourth, security, identity, and access management. The multi-cloud landscape complicates identity and access management. Different authentication models, the lack of a unified identity and access

management (IAM) system, and the emergence of «super identities» create critical vulnerabilities. A study by the scientific information portal «ResearchGate» identifies fragmented IAM systems as one of the key problems, along with disparate data classification [9].

Fifth, compliance and data sovereignty: In a multi-cloud environment, it is difficult to comply with various regulatory requirements such as GDPR, HIPAA, and local regulations.

Sixth, complexity of auditing, monitoring, and incident response. Multi-cloud architecture makes it difficult to collect, normalize, and correlate logs, as well as conduct audits. Without a unified SIEM architecture and normalized data, incident investigation processes and compliance are difficult.

Seventh, infrastructure complexity and operational costs. Data distribution and dependencies between cloud platforms increase traffic costs and require data duplication. The scientific portal ResearchGate points out the difficulty of abstracting security tools due to the heterogeneity of architectures, which increases costs and reduces efficiency [9]. Conducted analysis and practical experience of implementing data management (Data Governance) in multi-cloud environments reveal a number of typical barriers, which we have systematized in Table 1.

Table 1 - Main integration issues

No.	Problem	Manifestations in Multicloud
1	Fragmentation of platforms and policies	Different APIs, different policies, access conflicts
2	Shadow IT & data out of control	Unauthorized storage, risk of leaks, loss of quality
3	Lack of unity lineage	Difficulties with data tracing, auditing, quality
4	Fragmented IAM systems	Excessive privileges, undefined roles, increased attack surface
5	Regulatory inconsistency	Sovereignty, local norms, multi-jurisdictional control
6	Audit and log management	Uncoordinated logs , complex collection, incident management
7	Operating costs and expenses	Egress - boards, duplication, convergence of tools

Integration of data management practices (Data Governance) in a multi-cloud infrastructure requires not only an understanding of existing problems, but also the development of clear mechanisms to overcome them. Below are the main strategic directions and methods based on the analysis of current scientific and practical materials.

1. «Data Fabric» and «Data Mesh» architectures. «Data Fabric» is a centralized architectural layer that provides unified access to data from heterogeneous sources without the need to physically move it. This architecture uses metadata, semantic graphs, and automation to ensure data consistency and governance across the enterprise [6].

Data Mesh offers a decentralized approach where data management is delegated to domains – «data products». At the same time, a common platform provides standards and automated tools. This approach reduces fragmentation and accelerates the implementation of governance processes [10].

2. Policy -as- Code . The Policy-as-Code (PaC) concept allows security and compliance policies to be described as code. General-purpose engines such as Open Policy Agent (OPA), provide the ability to

centrally create declarative rules (in the Rego language) and automatically apply them to all cloud environments via CI/CD processes. This ensures uniformity, auditability, and automation of compliance processes [11]. PaC practices include storing policies in version control systems (e.g., Git) and integrating them with CI/CD pipelines for continuous monitoring and improvement.

3. Platforms for federated management. Platforms like Fybrik provide an open architecture for secure and controlled use of data in hybrid and multi-cloud environments. Such solutions support automatic policy enforcement, ensure privacy, and allow trade-offs between performance and compliance requirements. They also integrate with orchestration systems such as Kubernetes [12].

4. Simplify tools and practices. Using provider-neutral tools such as Atlan allows for the unification of processes such as data provenance tracking (data lineage), metadata management and tagging. This helps organizations avoid being dependent on vendor -specific tools (vendor lock-in) at the data management level.

Using resource abstraction and unified patterns, such as a single model for describing entities (e.g. «owner» and «cost-center» tags) and using a common CI/CD process, makes implementing governance much easier.

In response to identified data management integration issues (Data Governance) in multi-cloud architectures, a number of promising solutions are proposed by the scientific community and industrial practices. These approaches range from policy automation through Policy-as-Code to the implementation of unified metadata standards and the use of artificial intelligence (AI)-based tools for data discovery and classification. Below is a table that systematizes the key solutions, indicating their advantages and potential limitations.

Table 2 - Prospective Data Solutions Governance in Multicloud

No.	Approach / Architecture	Description and benefits
1	Data Fabric	Unify access and real-time control through metadata and automation
2	Data Mesh	Decentralization of governance through domain commands on a common platform
3	Policy-as-Code (OPA and etc.)	Code Policy Control Across All Clouds: Audit, Automation, Reliability
4	Federated Platforms (Fybrik)	Controlled access, privacy -oriented, support for hybrid environments
5	Independent Tools (Atlan)	Consolidated catalog, lineage and metadata governance
6	Abstraction and IaC patterns	Resource templates and policies built into CI/CD improve consistency

Data Management Integration (Data Governance) in multi-cloud environments is not only a technological but also a profound transformational challenge that requires taking into account the dynamics of architectural changes, regulatory frameworks and the emergence of new intelligent tools:

- management models based on artificial intelligence. Development of AI- Driven Governance involves the use of artificial intelligence for policy automation, cost optimization (FinOps) and adaptive security. For example, the model proposed by scientist Sunil Kattikar, includes cost

forecasting, dynamic policy enforcement and monitoring through AI dashboards [13]. This allows organizations to adapt to changing conditions without constant manual intervention;
- architectures based on the Zero principle Trust. Transition to Zero Architecture Trust is a key step in maintaining security in multi-cloud environments. As scientific research shows , this approach helps to eliminate risks associated with excessive access rights and distributed identities, which is especially relevant for hybrid and multi-cloud infrastructures [14];
- cryptographic methods and trusted intermediaries. New research suggests the use of advanced cryptographic methods such as «Secure Multi-Party Computation» (MPC) and «Fully Homomorphic Encryption» (FHE). These technologies can become the basis for creating trusted "trustless" intermediaries within data spaces (Data Spaces), which will improve privacy and trust standards in multi-cloud ecosystems;
Data evolution Mesh and federated governance. Data Mesh is transforming from a purely architectural solution into a complex organizational and technological platform, which requires further academic understanding;
- convergence of multi-cloud and edge architectures. The future of multi-cloud environments is seen in automation, AI-optimization of resources and their tight integration with edge computing;
- digital sovereignty and intermediate platforms. The trend is «cloud repatriation», or the return of workloads to private or hybrid environments, driven by the desire for greater control, resilience, and regulatory compliance. This trend underscores the importance of a hybrid strategy and federated governance. In addition, initiatives such as Gaia -X in Europe aim to create a cloud infrastructure focused on ensuring digital sovereignty and reducing dependence on large technology companies [15].

Table 3 - Areas for Future Research in Governance multiclouds

No	**Field of study**	**Possible direction/call**
1	AI-driven Governance	Application of FinOps, predictive policies, adaptation of Security as code
2	Zero Trust and Privileged Architecture	Identity and Least Rights Management in Multi-Cloud Beaches
3	Cryptographic methods: MPC, FHE, Data Spaces	Secure exchange and privacy between environments
4	Data Architectures Mesh	Self-service platforms, federated management and data management as a product
5	Edge + Multi-Cloud Integration	Connecting Clouds and Edge Devices with Automation and AI Optimization
6	Sovereignty and regional ecosystems	Cloud repatriation, Gaia -X, hybrid models with transparent economy and compliance

Data Management Integration (Data Governance) with multi-cloud platforms is one of the key challenges for modern organizations seeking to take advantage of the flexibility and scalability of cloud technologies. The challenges associated with platform fragmentation, security, metadata management and organizational structure are not insurmountable, but require a systematic approach. Successful integration is possible with the implementation of a balanced strategy that includes: unification of metadata and the use of Data architectures Fabric or Data Mesh ; automation of security policies (e.g.,

using Policy-as-Code); implementation of centralized key management systems; creation of an effective organizational structure for data management.

This approach will ensure data integrity, security and compliance with regulatory requirements in a complex and distributed multi-cloud infrastructure.

References

1. ISO/IEC 22123-1:2021. Information technology - Cloud computing - Concepts and terminology - Part 1: Vocabulary.

2. Cyfuture . Vendor lock-in & more: why multi-cloud solutions matter [Electronic resource] // LinkedIn, 2024. - Mode access : https://www.linkedin.com/pulse/vendor-lock-in-more-why-multi-cloud-solutions-matter-cyfuture-ptqzc (date accesses: 13.08.2025).

3. Reece T ., Ali S. , Abualhaol I ., Musa S. , Alenezi M . Cybersecurity risks in multi - cloud environments [Electronic resource] // arXiv preprint , 2023. - Mode access : https :// arxiv . org / abs / 2306.01862 (date accessed: 13.08.2025).

4. Fernandes DAB, Soares LFB, Gomes JV, Freire MM, Inácio PRM Security issues in cloud environments: a survey // International Journal of Information Security. - 2014. - T. 13, No. 2. - pp. 113–170 . - DOI: 10.1007/s10207-013-0208-7.

5. Phaphom N., Abrahamsson A., Hakansson L., Paweenawat P.V., Kihl M. A survey study on major technical barriers affecting the decision to adopt cloud services [Electronic resource] // arXiv preprint, 2017. - Mode access: https://arxiv.org/abs/1712.00606 (date accessed: 14.08.2025).

6. Springer. Cloud computing - concepts and technologies [Electronic resource]. - Mode access : https://link.springer.com/chapter/10.1007/978-3-319-94472-2_1 (date accessed : 13.08.2025).

7. Folio3 Cloud Services. Cloud data governance [Electronic resource]. - Mode access : https :// cloud . folio 3. com / blog / cloud - data - governance (date accessed: 12.08.2025).

8. Maruti Techlabs . Best practices: data lineage in multi-cloud [Electronic resource]. - Mode Access : https :// marutitech . com / best - practices - data - lineage - multi - cloud (date accessed: 12.08.2025).

9. ResearchGate . Security and compliance challenges in multi-cloud environments: a comparative study of industry-specific strategies [Electronic resource]. - Mode access : https://www.researchgate.net/publication/391442348_Security_and_Compliance_Challenges_in_Multi-Cloud_Environments_A_Comparative_Study_of_Industry-Specific_Strategies (date accesses: 14.08.2025).

10. Wikipedia . Data mesh [Electronic resource]. - Access mode: https :// en . wikipedia . org / wiki / Data _ mesh (date of access: 12.08.2025).

11. Uplatz . Policy-as-code for multi-cloud security governance [Electronic resource]. - Mode access : https://uplatz.com/blog/policy-as-code-for-multi-cloud-security-governance (date accessed: 13.08.2025).

12. MDPI . Future Internet : Vol . 16, Issue 4 [Electronic resource]. - Access mode: https :// www . mdpi . com /1999-5903/16/4/115 (date of access: 14.08.2025).

13. IJERT . AI - driven hybrid multi - cloud governance strategy [Electronic resource]. - Access mode: https :// www . ijert . org / ai - driven - hybrid - multi - cloud - governance - strategy (date of access: 14.08.2025).

14. IAEME . Multi - cloud data governance approaches [Electronic resource]. - Access mode: https :// iaeme . com / Home / article _ id / IJRCAIT _08_01_035 (date of access: 12.08.2025).
15. Wired . Europe 's Gaia - X cloud project [Electronic resource]. - Access mode: https :// www . wired . com / story / europe - gaia - x - cloud - amazon - google (date of access: 13.08.2025).