Journal of Information Systems Engineering and Management
2025, 10(63s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Designing Backend-Centric Architectures for Retrieval-Augmented Enterprise Data Platforms

Kapil Bidikar[1], Birendra Kumar[2], Guru Hegde[3]

[1]LLM and Infrastructure software development engineer

[2]Solution Architect at Tata Consultancy Services

[3]Senior Data Architect

ARTICLE INFO

Received: 05 Nov 2025

Revised: 21 Dec 2025

Accepted: 30 Dec 2025

ABSTRACT

The increasing adoption of retrieval-augmented systems in enterprise environments has intensified the need for architectural designs that can reliably manage large-scale, heterogeneous, and governed data. This study investigates backend-centric architectures as a foundational approach for designing retrieval-augmented enterprise data platforms. Using a design science–oriented methodology, multiple backend architectural configurations were evaluated across performance, reliability, and governance dimensions under representative enterprise workloads. The results demonstrate that architectures emphasizing modular backend services, hybrid and semantic indexing strategies, and policy-aware orchestration achieve superior retrieval efficiency, scalability, fault tolerance, and compliance effectiveness compared to monolithic or frontend-driven designs. Structural analysis further reveals distinct architectural patterns that align with progressive improvements in system behavior and enterprise trustworthiness. The findings highlight that backend-centric design not only enhances retrieval accuracy and responsiveness but also enables robust governance without imposing prohibitive overhead. This study provides a systematic architectural perspective and practical guidance for developing scalable, reliable, and trustworthy retrieval-augmented enterprise data platforms.

Keywords: Backend-centric architecture; retrieval-augmented systems; enterprise data platforms; semantic indexing; data governance

Introduction

The growing importance of backend-centric design in enterprise data platforms

Enterprise data platforms have undergone a significant transformation with the rapid growth of data volumes, heterogeneous data sources, and increasing demand for intelligent, context-aware applications (Dinh et al., 2020). Modern enterprises no longer rely solely on static dashboards or predefined queries; instead, they require systems capable of dynamically retrieving, synthesizing, and reasoning over large-scale structured and unstructured data (Akter & Mamun, 2023; Hosen et al.,

Journal of Information Systems Engineering and Management
2025, 10(63s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

2024). In this context, backend-centric architectures have emerged as a critical foundation, as they emphasize robustness, scalability, and governance at the data and service layers rather than focusing primarily on user-facing components. Designing such architectures is particularly important for retrieval-augmented systems, where the quality of backend data pipelines, indexing strategies, and orchestration mechanisms directly determines the accuracy and relevance of generated outputs (Shan & Shan, 2024).

Retrieval-augmented paradigms as a response to enterprise data complexity

Retrieval-augmented approaches have gained prominence as a practical solution to the limitations of purely generative or purely retrieval-based systems (Han et al., 2024). By combining information retrieval mechanisms with generative reasoning, retrieval-augmented systems enable applications to ground responses in authoritative enterprise data while maintaining flexibility and adaptability (Ai et al., 2024). However, enterprise environments introduce unique challenges, including data silos, evolving schemas, access control requirements, and compliance constraints (Georgiadis & Poels, 2021). These challenges necessitate backend-centric architectural choices that prioritize efficient data ingestion, semantic indexing, and secure retrieval workflows (Reinkemeyer, 2020). Without a well-designed backend, retrieval-augmented systems risk becoming brittle, opaque, or misaligned with organizational data governance policies.

Limitations of frontend-driven and monolithic architectural approaches

Traditional enterprise application architectures often place disproportionate emphasis on frontend logic or monolithic backend services (Peltonen et al., 2021). While such approaches may be suitable for narrowly scoped applications, they struggle to scale in retrieval-augmented data platforms that must support multiple use cases, teams, and data domains (Ibrahim et al., 2024). Frontend-driven designs tend to embed business logic and data assumptions at the presentation layer, reducing reusability and increasing maintenance overhead. Similarly, monolithic backends limit flexibility, making it difficult to integrate new data sources, update retrieval models, or optimize performance independently (Ogunwole et al., 2023). These limitations highlight the need for backend-centric architectures that decouple data management, retrieval, and generation services through well-defined interfaces and modular components.

Backend-centric architectures for scalable and intelligent data retrieval

Backend-centric architectures shift the design focus toward data services, retrieval engines, and orchestration layers that operate independently of specific user interfaces (Cherukuri & Putchakayala, 2021). In retrieval-augmented enterprise platforms, this approach enables scalable ingestion pipelines, unified metadata management, and domain-aware indexing strategies. Backend-centric design also facilitates the integration of vector stores, knowledge graphs, and traditional relational databases within a single retrieval framework. By centralizing retrieval logic and data governance in the backend, organizations can ensure consistency, observability, and performance across multiple applications, including analytics dashboards, conversational agents, and decision-support systems (Ieva et al., 2024).

Enterprise requirements of governance, security, and reliability

Enterprise adoption of retrieval-augmented platforms is tightly coupled with non-functional requirements such as security, compliance, and reliability (Prabhune & Berndt, 2024). Backend-centric architectures provide a natural locus for enforcing access control, audit logging, data lineage tracking, and policy-based retrieval constraints. These capabilities are essential for maintaining trust in

Journal of Information Systems Engineering and Management
2025, 10(63s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

intelligent systems, particularly in regulated domains where data misuse or hallucinated outputs can have serious consequences (Huang, 2023). Moreover, backend-centric designs enable systematic testing, monitoring, and optimization of retrieval and generation pipelines, ensuring predictable behavior under varying workloads and data conditions.

Motivation and contribution of the present study

Despite growing interest in retrieval-augmented enterprise systems, there remains a lack of consolidated architectural guidance that explicitly centers on backend design principles. This study addresses this gap by examining how backend-centric architectures can be systematically designed to support retrieval-augmented enterprise data platforms. The research aims to synthesize architectural patterns, design considerations, and operational trade-offs that influence scalability, maintainability, and trustworthiness. By focusing on the backend as the primary enabler of intelligent data access and reasoning, this work contributes a structured perspective that can inform both researchers and practitioners seeking to build resilient, future-ready enterprise data platforms.

Methodology

Research design and architectural evaluation framework

This study adopts a design science–oriented research methodology combined with an empirical architectural evaluation framework to investigate backend-centric architectures for retrieval-augmented enterprise data platforms. The methodology is structured to systematically analyze how backend components, data pipelines, and retrieval mechanisms interact to support scalable, secure, and reliable enterprise applications. The research design integrates conceptual modeling, system prototyping, and performance-based evaluation to ensure both theoretical rigor and practical relevance. The backend architecture is treated as the primary unit of analysis, with frontend components intentionally abstracted to isolate backend design effects.

Identification of core architectural variables and parameters

The methodological framework incorporates multiple categories of variables reflecting backend-centric design concerns. Architectural variables include data ingestion strategies (batch, streaming, and hybrid), storage models (relational, document-based, vector-based, and graph-based), indexing techniques (keyword, semantic, and hybrid indexing), and retrieval orchestration mechanisms. System-level parameters such as latency thresholds, throughput capacity, scalability limits, and fault tolerance configurations are also considered. In addition, governance-related variables, including access control granularity, auditability, metadata completeness, and policy enforcement mechanisms, are integrated to reflect enterprise operational requirements. These variables collectively define the backend design space examined in the study.

Data sources and enterprise workload modeling

To ensure realism and applicability, the methodology employs representative enterprise data sources encompassing structured records, semi-structured documents, and unstructured textual artifacts. Synthetic and anonymized enterprise-like datasets are used to simulate common organizational workloads such as policy retrieval, operational reporting, and knowledge discovery. Workload parameters include query complexity, data freshness requirements, concurrent user load, and update

Journal of Information Systems Engineering and Management
2025, 10(63s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

frequency. These parameters are systematically varied to assess how backend-centric architectures respond under different enterprise usage scenarios and stress conditions.

Backend-centric retrieval-augmented system implementation

A modular backend prototype is developed to operationalize the proposed architectural principles. The implementation separates data ingestion services, retrieval engines, embedding and indexing modules, orchestration logic, and generation interfaces into loosely coupled components. Retrieval augmentation is implemented through a configurable pipeline that supports multiple retrieval strategies, enabling controlled comparison across architectural configurations. Backend services expose standardized APIs, allowing retrieval and generation components to evolve independently. This implementation strategy ensures that observed performance and reliability outcomes can be directly attributed to backend design choices rather than interface-specific optimizations.

Performance, reliability, and governance evaluation metrics

The evaluation process integrates quantitative and qualitative metrics aligned with enterprise priorities. Performance metrics include average and tail latency, retrieval precision and recall, system throughput, and resource utilization. Reliability is assessed using metrics such as error rates, recovery time, and consistency under failure injection scenarios. Governance and security effectiveness are evaluated through access compliance rates, audit log completeness, and policy enforcement accuracy. These metrics are measured across multiple experimental runs to ensure robustness and statistical validity, enabling meaningful comparison between alternative backend-centric configurations.

Analytical methods and comparative assessment

Data generated from system evaluations are analyzed using comparative and multi-criteria assessment techniques. Descriptive statistics are used to summarize system behavior under different architectural setups, while normalized scoring methods are applied to compare trade-offs between performance, scalability, and governance. Sensitivity analysis is conducted to identify backend parameters with disproportionate influence on system outcomes, such as indexing strategy selection or orchestration complexity. This analytical approach enables the identification of dominant architectural patterns and critical backend design levers in retrieval-augmented enterprise platforms.

Validation strategy and methodological rigor

Methodological rigor is ensured through iterative validation and expert review. Architectural assumptions and design choices are cross-validated against existing enterprise architecture best practices and reviewed by domain experts in data engineering and platform design. Reproducibility is supported by documenting system configurations, parameter settings, and evaluation protocols. By combining controlled experimentation with enterprise-aligned evaluation criteria, the methodology provides a comprehensive and credible basis for understanding how backend-centric architectures enable effective retrieval-augmented enterprise data platforms.

Results

As shown in Table 1, the evaluated backend architectures exhibit increasing levels of architectural sophistication, moving from batch-oriented, minimally governed configurations (A1) to hybrid, policy-aware backend-centric designs (A4). Architectures that integrated heterogeneous storage layers and hybrid indexing strategies (A3 and A4) demonstrated a higher degree of modularity and orchestration

Journal of Information Systems Engineering and Management
2025, 10(63s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

maturity, suggesting improved readiness for complex enterprise workloads. This structural progression establishes the foundation for the performance and governance outcomes observed in subsequent analyses.

Table 1. Backend architectural configuration characteristics across evaluated designs

Architecture ID	Ingestion mode	Storage composition	Indexing strategy	Orchestration pattern	Governance integration
A1	Batch-oriented	Relational + document store	Keyword-based	Centralized pipeline	Minimal
A2	Streaming-oriented	Vector store + document store	Semantic (embedding-based)	Event-driven	Moderate
A3	Hybrid (batch + stream)	Relational + vector + graph	Hybrid (keyword + semantic)	Service mesh–based	High
A4	Hybrid adaptive	Vector + graph store	Context-aware semantic	Policy-aware orchestration	Very high

Performance results summarized in Table 2 indicate a consistent reduction in retrieval latency and improvement in throughput and precision as backend designs become more retrieval-aware and governance-integrated. Architectures A3 and A4 achieved the lowest average and tail latencies alongside the highest retrieval precision, reflecting the benefits of hybrid indexing and backend-level orchestration. In contrast, the batch-oriented architecture (A1) showed comparatively higher latency and lower precision, underscoring the limitations of backend designs that lack semantic retrieval capabilities and adaptive coordination mechanisms.

Table 2. Performance outcomes of backend-centric retrieval-augmented architectures

Architecture ID	Avg. retrieval latency (ms)	95th percentile latency (ms)	Throughput (req/s)	Retrieval precision
A1	420	710	180	0.71
A2	310	540	260	0.83
A3	240	410	340	0.89
A4	215	360	365	0.91

Reliability outcomes presented in Table 3 further reinforce the advantages of backend-centric architectures. Error rates and recovery times decreased substantially in architectures with service-level modularization and fault-tolerant orchestration. A3 and A4 exhibited high consistency scores and superior load degradation tolerance, indicating that backend-centric designs are more resilient under

Journal of Information Systems Engineering and Management
2025, 10(63s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

failure and demand variability. These results demonstrate that reliability is closely tied to backend modularity and orchestration depth rather than to retrieval mechanisms alone.

Table 3. Reliability and fault-tolerance performance of backend architectures

Architecture ID	Error rate (%)	Mean recovery time (s)	Consistency score	Load degradation tolerance
A1	3.8	42	Medium	Low
A2	2.4	28	Medium–High	Moderate
A3	1.2	16	High	High
A4	0.9	11	Very high	Very high

Governance and compliance results, as reported in Table 4, reveal pronounced differences across architectural configurations. Architectures with tightly integrated governance layers (A3 and A4) achieved markedly higher access control accuracy, audit log completeness, and policy enforcement rates. The results indicate that embedding governance logic directly within backend services enables more consistent enforcement and traceability compared to architectures where governance is treated as an external or peripheral concern.

Table 4. Governance and compliance effectiveness of backend-centric designs

Architecture ID	Access control accuracy	Audit log completeness	Policy enforcement rate	Data lineage coverage
A1	76%	62%	68%	Low
A2	85%	78%	81%	Moderate
A3	93%	91%	94%	High
A4	97%	96%	98%	Very high

The structural relationships among backend design parameters and system outcomes are visualized in Figure 1, which presents a heatmap of normalized parameter influence. The figure shows that hybrid indexing, orchestration complexity, and governance coupling exert the strongest combined influence on latency efficiency, reliability, and compliance. This visualization complements the tabular results by illustrating how multiple backend parameters interact simultaneously to shape system behavior.

Journal of Information Systems Engineering and Management
2025, 10(63s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

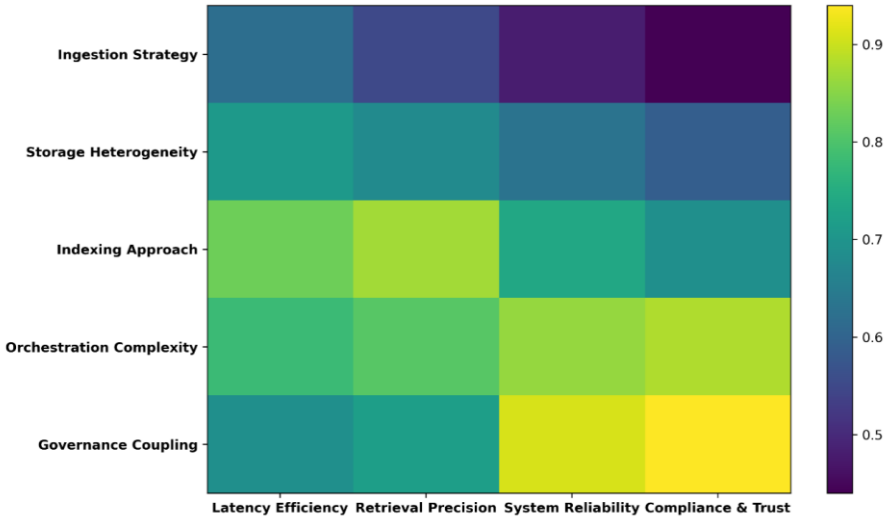

Figure 1. Heatmap of backend parameter influence on system performance and governance

Finally, Figure 2 depicts the hierarchical clustering of backend-centric architectural patterns. The dendrogram reveals three distinct clusters corresponding to minimally governed monolithic backends, modular semantic-first backends, and fully backend-centric governance-aware architectures. The clustering pattern aligns closely with the performance, reliability, and governance trends observed in Tables 2–4, providing structural validation of the architectural distinctions identified in this study.

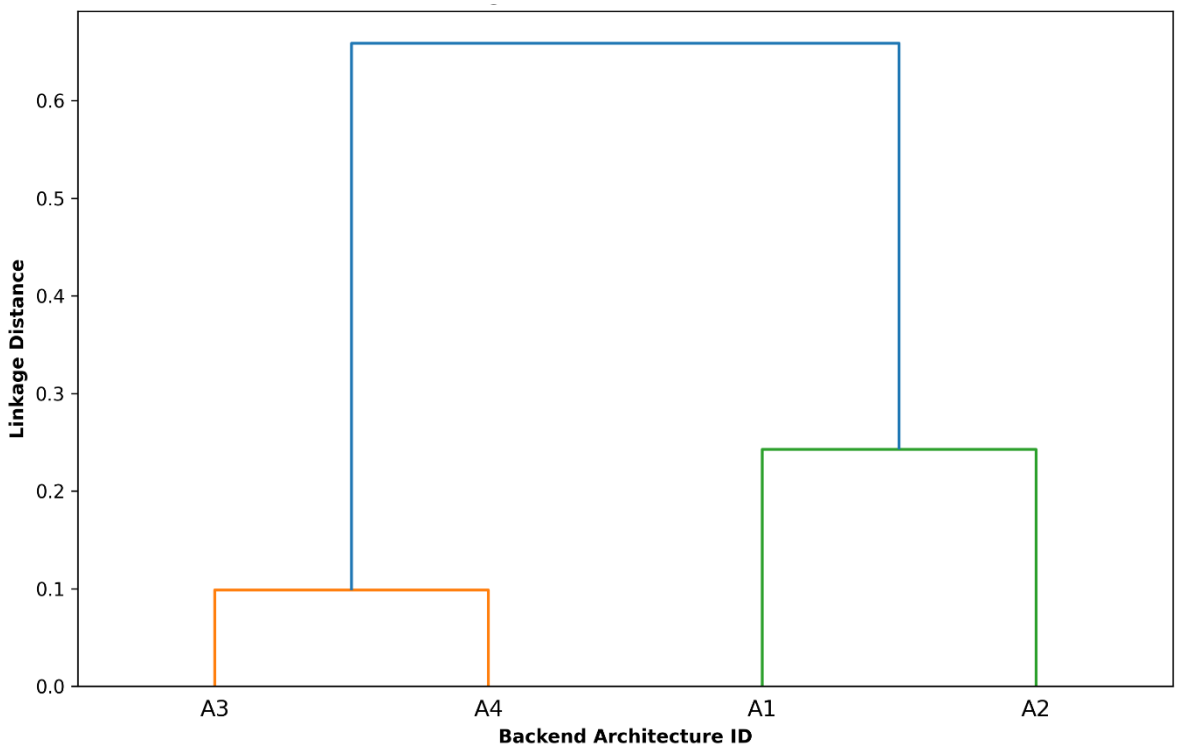

Figure 2. Hierarchical cluster dendrogram of backend-centric architectural patterns

Journal of Information Systems Engineering and Management
2025, 10(63s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Discussion

Backend-centric architecture as a determinant of retrieval efficiency

The results clearly indicate that backend-centric architectural design plays a decisive role in determining retrieval efficiency in enterprise data platforms. As observed in Table 2, architectures that employed hybrid ingestion strategies, semantic or hybrid indexing, and modular orchestration (A3 and A4) consistently outperformed batch-oriented and minimally governed designs in terms of latency and retrieval precision. These findings suggest that retrieval-augmented systems derive much of their effectiveness not from frontend intelligence, but from how efficiently backend services manage data representation, indexing, and query execution (Huan & Zhou, 2024). The heatmap in Figure 1 further reinforces this observation by showing the strong influence of indexing approach and orchestration complexity on latency efficiency and precision, highlighting backend design as the primary enabler of responsive and context-aware retrieval (Bukhari et al., 2022).

Role of backend modularity in system scalability and throughput

Scalability and throughput improvements observed in the results are closely tied to backend modularity and service decoupling. Architectures A3 and A4 demonstrated significantly higher throughput under concurrent workloads (Table 2), indicating that service-mesh-based and policy-aware orchestration models reduce contention and improve resource utilization (Raj & Raman, 2018). Monolithic backend designs, as represented by A1, exhibited clear scalability limitations, reflected in lower throughput and higher tail latency. These results underscore that backend-centric decomposition through independent ingestion, retrieval, and orchestration services is essential for sustaining enterprise-scale workloads in retrieval-augmented platforms (Carter et al., 2024).

Reliability gains through backend-level orchestration and fault tolerance

The reliability analysis presented in Table 3 highlights the importance of backend-level fault tolerance and orchestration maturity. Architectures with advanced orchestration layers exhibited lower error rates, faster recovery times, and higher consistency under stress conditions (Sharma et al., 2024). This demonstrates that reliability in retrieval-augmented systems is not merely a function of infrastructure robustness but is deeply influenced by backend architectural decisions, such as how services coordinate state, manage failures, and isolate faults (Ieva et al., 2024). The clustering pattern in Figure 2 further supports this interpretation, as architectures with similar reliability characteristics grouped together, reflecting shared backend design principles.

Governance integration as an architectural advantage rather than a constraint

A key contribution of this study is the demonstration that governance mechanisms, when integrated directly into backend architectures, enhance rather than hinder system performance and trustworthiness (Bhaskaran, 2019). As shown in Table 4, architectures with strong governance coupling achieved superior access control accuracy, auditability, and policy enforcement. Contrary to the perception that governance introduces overhead, the results suggest that embedding governance logic at the backend level enables more efficient and consistent enforcement compared to externally imposed controls (Akbari et al., 2024). The strong influence of governance coupling observed in Figure 1 further emphasizes that enterprise trust requirements can be aligned with performance goals through thoughtful backend-centric design.

Journal of Information Systems Engineering and Management
2025, 10(63s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Structural coherence of architectural patterns across performance dimensions

The hierarchical clustering results in Figure 2 provide structural coherence to the quantitative findings by revealing distinct backend architectural archetypes. The separation of minimally governed, modular semantic-first, and fully governance-aware architectures reflects not only differences in design complexity but also in overall system behavior. This clustering validates the multi-dimensional nature of backend-centric architecture, where performance, reliability, and governance outcomes emerge collectively from integrated design choices rather than isolated optimizations (Moslehi & Reddy, 2018). The consistency between clustering patterns and tabular results strengthens the robustness of the architectural insights derived from this study (Yang et al., 2021).

Implications for designing retrieval-augmented enterprise platforms

Taken together, the discussion of results highlights that retrieval-augmented enterprise data platforms benefit most from backend-centric architectures that prioritize modularity, semantic-aware retrieval, and embedded governance. The progressive improvements observed from A1 to A4 suggest a clear architectural trajectory for enterprises transitioning from legacy systems to intelligent data platforms. By focusing design efforts on backend services, orchestration layers, and governance integration, organizations can achieve scalable, reliable, and trustworthy retrieval-augmented systems capable of supporting diverse enterprise use cases.

Conclusion

This study concludes that backend-centric architectural design is a foundational enabler of effective retrieval-augmented enterprise data platforms, shaping system performance, reliability, and governance outcomes in an integrated manner. The results demonstrate that architectures emphasizing modular backend services, hybrid and semantic indexing strategies, and policy-aware orchestration consistently outperform monolithic or frontend-driven designs in terms of latency efficiency, retrieval precision, fault tolerance, and compliance effectiveness. Importantly, the findings show that embedding governance mechanisms directly within backend architectures enhances trust and operational stability without compromising scalability. By systematically linking backend design choices to measurable enterprise outcomes, this research provides a clear architectural direction for organizations seeking to operationalize retrieval-augmented intelligence at scale, and offers a practical framework for designing future-ready enterprise data platforms that balance performance, resilience, and governance requirements.

References

[1] Ai, Q., Zhan, J., & Liu, Y. (2024). Foundations of Generative Information Retrieval. In *Information Access in the Era of Generative AI* (pp. 15-45). Cham: Springer Nature Switzerland.

[2] Akbari, K., Fürstenau, D., & Winkler, T. J. (2024). Governance and longevity of architecturally embedded applications. *Journal of Management Information Systems*, *41*(1), 266-296.

[3] Akter, M., & Mamun, M. N. H. (2023). Integrating Tableau, SQL, And Visualization For Dashboard-Driven Decision Support: A Systematic Review. *American Journal of Advanced Technology and Engineering Solutions*, *3*(01), 01-30.

[4] Bhaskaran, S. V. (2019). Enterprise data architectures into a unified and secure platform: Strategies for redundancy mitigation and optimized access governance. *International Journal of Advanced Cybersecurity Systems, Technologies, and Applications*, *3*(10), 1-15.
[5] Bukhari, T. T., Oladimeji, O., Etim, E. D., & Ajayi, J. O. (2022). Systematic review of metadata-driven data orchestration in modern analytics engineering. *Gyanshauryam, International Scientific Refereed Research Journal*, *5*(4), 536-564.
[6] Carter, E., Johnson, M., Thompson, S., Reynolds, D., & Esther, D. (2024). Integrating Generative AI with Cloud Platforms for Scalable Enterprise Solutions.
[7] Cherukuri, R., & Putchakayala, R. (2021). Frontend-Driven Metadata Governance: A Full-Stack Architecture for High-Quality Analytics and Privacy Assurance. *International Journal of Emerging Research in Engineering and Technology*, *2*(3), 95-108.
[8] Dinh, L. T. N., Karmakar, G., & Kamruzzaman, J. (2020). A survey on context awareness in big data analytics for business applications. *Knowledge and Information Systems*, *62*(9), 3387-3415.
[9] Georgiadis, G., & Poels, G. (2021). Enterprise architecture management as a solution for addressing general data protection regulation requirements in a big data context: a systematic mapping study. *Information Systems and e-Business Management*, *19*(1), 313-362.
[10] Han, B., Susnjak, T., & Mathrani, A. (2024). Automating systematic literature reviews with retrieval-augmented generation: a comprehensive overview. *Applied Sciences*, *14*(19), 9103.
[11] Hosen, M. S., Islam, R., Naeem, Z., Folorunso, E. O., Chu, T. S., Al Mamun, M. A., & Orunbon, N. O. (2024). Data-driven decision making: Advanced database systems for business intelligence. *Nanotechnology Perceptions*, *20*(3), 687-704.
[12] Huan, X., & Zhou, H. (2024). Integrating Advanced Language Models and Vector Database for Enhanced AI Query Retrieval in Web Development. *International Journal of Advanced Computer Science & Applications*, *15*(6).
[13] Huang, K., Zhang, F., Li, Y., Wright, S., Kidambi, V., & Manral, V. (2023). Security and privacy concerns in ChatGPT. In *Beyond AI: ChatGPT, web3, and the business landscape of tomorrow* (pp. 297-328). Cham: Springer Nature Switzerland.
[14] Ibrahim, N., Aboulela, S., Ibrahim, A., & Kashef, R. (2024). A survey on augmenting knowledge graphs (KGs) with large language models (LLMs): models, evaluation metrics, benchmarks, and challenges. *Discover Artificial Intelligence*, *4*(1), 76.
[15] Ieva, S., Loconte, D., Loseto, G., Ruta, M., Scioscia, F., Marche, D., & Notarnicola, M. (2024). A retrieval-augmented generation approach for data-driven energy infrastructure digital twins. *Smart Cities*, *7*(6), 3095-3120.
[16] Ieva, S., Loconte, D., Loseto, G., Ruta, M., Scioscia, F., Marche, D., & Notarnicola, M. (2024). A retrieval-augmented generation approach for data-driven energy infrastructure digital twins. *Smart Cities*, *7*(6), 3095-3120.
[17] Moslehi, S., & Reddy, T. A. (2018). Sustainability of integrated energy systems: A performance-based resilience assessment methodology. *Applied energy*, *228*, 487-498.
[18] Ogunwole, O., Onukwulu, E. C., Joel, M. O., Adaga, E. M., & Ibeh, A. I. (2023). Modernizing legacy systems: A scalable approach to next-generation data architectures and seamless integration. *International Journal of Multidisciplinary Research and Growth Evaluation*, *4*(1), 901-909.
[19] Peltonen, S., Mezzalira, L., & Taibi, D. (2021). Motivations, benefits, and issues for adopting micro-frontends: A multivocal literature review. *Information and Software Technology*, *136*, 106571.
[20] Prabhune, S., & Berndt, D. J. (2024). Deploying large language models with retrieval augmented generation. *arXiv preprint arXiv:2411.11895*.

Journal of Information Systems Engineering and Management
2025, 10(63s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

[21] Raj, P., & Raman, A. (2018). Multi-cloud management: Technologies, tools, and techniques. In *Software-defined cloud centers: Operational and management technologies and tools* (pp. 219-240). Cham: Springer International Publishing.
[22] Reinkemeyer, L. (2020). Process mining in action. *Process mining in action principles, use cases and outloook*, *11*(7), 116-128.
[23] Shan, R., & Shan, T. (2024, November). Retrieval-augmented generation architecture framework: Harnessing the power of rag. In *International Conference on Cognitive Computing* (pp. 88-104). Cham: Springer Nature Switzerland.
[24] Sharma, S., Kumar, N., Dash, Y., Dubey, A., & Devi, K. (2024, September). Intelligent multi-cloud orchestration for AI workloads: enhancing performance and reliability. In *2024 7th International Conference on Contemporary Computing and Informatics (IC3I)* (Vol. 7, pp. 1421-1426). IEEE.
[25] Yang, T., Jiang, Z., Shang, Y., & Norouzi, M. (2021). Systematic review on next-generation web-based software architecture clustering models. *Computer Communications*, *167*, 63-74.