

A Comprehensive Review of the Impact of Snowflake AI Data Warehouse in the Data Analytics Domain

Prof. Shahin Shoukat Makubhai¹, Mayur Kankankwadi², Aditya Kulkarni³,
Mukund Kulkarni⁴ Aditya Patil⁵

Professor, Department of Computer Science and Engineering, MIT School of Computing, Pune, India ¹
Student, Department of Computer Science and Engineering, MIT School of Computing, Pune, India ^{2,3,4,5}

Abstract: *The rapid evolution of the digital landscape demands advanced data warehousing solutions capable of managing massive, diverse datasets while supporting high-performance business intelligence. This review comprehensively examines the transformative impact of the Snowflake cloud-native data platform on the data analytics and artificial intelligence (AI) domains.*

At the core of Snowflake's innovation is its multi-cluster shared-data architecture, which fundamentally decouples storage from compute layers, enabling independent, near-infinite scalability and cost-efficient resource management. We explore how this distinct architecture overcomes the limitations of traditional on-premise systems by effortlessly processing structured, semi-structured, and unstructured data natively without compromising performance.

Furthermore, this paper highlights Snowflake's profound integration with machine learning (ML) and AI workflows. Features such as Snowpark, Snowflake Cortex AI, and built-in ML functions empower data scientists to execute complex data transformations, feature engineering, and model inference directly within the data environment. This paradigm significantly reduces data movement and latency, while the platform's seamless interoperability with major cloud providers (AWS, Azure, GCP) and external AI frameworks (e.g., TensorFlow, PyTorch) fosters a highly flexible and collaborative analytics ecosystem. Additionally, we evaluate Snowflake's robust data governance, advanced security protocols, and zero-copy data sharing capabilities, which are essential for maintaining regulatory compliance (e.g., HIPAA, GDPR) and securing sensitive information in highly regulated industries like healthcare and finance. By analyzing real-world use cases, predictive analytics optimization, and platform scaling methodologies, this review demonstrates that Snowflake not only democratizes data access across organizations but also accelerates the transition from traditional, reactive business intelligence to proactive, real-time, AI-driven decision-making. Ultimately, the convergence of cloud computing, AI/ML, and Snowflake's optimized architecture establishes a future-ready, resilient foundation for modern enterprise data strategies.

Keywords: Snowflake, Data warehouse, Data Analytics, Lakehouse Architecture, OLAP, Cloud, AI

I. INTRODUCTION

Background on Data Warehousing In the modern digital era, data has emerged as a paramount resource for organizations to drive strategic decisions, enhance operational efficiency, and gain a competitive advantage. To harness this vast amount of information, data warehousing solutions serve as centralized repositories that aggregate and organize data from multiple sources, forming the backbone for business intelligence (BI) and advanced analytics [11]. As the volume, velocity, and variety of data continue to grow exponentially, the need for robust, highly scalable, and cost-effective data management platforms has never been more critical [3].



Challenges of Traditional Data Warehouses Historically, enterprises relied on traditional, on-premises data warehousing solutions that are increasingly struggling to meet the demands of modern analytics [3][11]. These legacy systems were designed for small, static clusters and predictable workloads, tightly coupling storage and compute resources. Consequently, they suffer from restricted scalability, sluggish query performance, data silos, and high maintenance costs. When faced with heterogeneous workloads or the need to scale, traditional systems require expensive hardware upgrades and complex data reshuffling. Furthermore, they heavily depend on rigid Extract, Transform, Load (ETL) pipelines that are ill-equipped to handle the rapid influx of schema-less, semi-structured, and unstructured data generated by modern applications, IoT devices, and social media.

Snowflake Innovations To address these fundamental shortcomings, Snowflake was developed as a pioneering, fully cloud-native data platform [14]. At the heart of Snowflake's innovation is its unique multi-cluster shared-data architecture, which completely decouples the storage layer from the compute layer [14][4]. This design allows organizations to elastically and independently scale computing resources on demand without affecting data integrity or experiencing downtime. Snowflake fundamentally transforms data accessibility by natively supporting structured, semi-structured (such as JSON, Avro, and Parquet), and unstructured data within a single platform [3][11]. Additionally, it introduces revolutionary features like Secure Data Sharing, which allows instantaneous collaboration across departments and external partners without copying or moving data, and Zero-Copy Cloning, which enables the instant creation of database snapshots for development and testing without incurring additional storage costs [5].

Analytics and AI Impact The advent of Snowflake has profoundly impacted the landscape of big data management and artificial intelligence (AI). By providing a highly scalable and secure environment, Snowflake eliminates the data movement friction that traditionally hampered machine learning (ML) workflows. Through innovations like Snowpark, data scientists and engineers can execute complex data transformations, feature engineering, and ML model training using familiar languages (Python, Java, Scala) directly within the Snowflake compute engine [8]. Furthermore, its seamless integration with external AI frameworks (such as TensorFlow, PyTorch, and Amazon SageMaker) and built-in ML functions empower organizations to shift from reactive historical reporting to proactive, predictive, and real-time decision-making [8].

Aim of the Paper The primary aim of this review paper is to provide a comprehensive evaluation of Snowflake's architecture and its transformative impact on the data analytics and machine learning domains. By examining its core technical components, scalability, and integration capabilities, this study seeks to illuminate how Snowflake overcomes the limitations of traditional data warehouses. Furthermore, this paper will explore real-world use cases, performance optimization strategies, and future trends—such as the integration of generative AI and Domain-AI agents—to guide decision-makers in adopting robust, future-ready data strategies.

II. LITERATURE REVIEW

1. Evolution from Traditional Data Warehouses to Cloud-Native Platforms

The rapid digitalization of modern enterprises has driven an exponential increase in the volume, velocity, and variety of organizational data, necessitating a fundamental shift in data management strategies. Historically, enterprises relied on traditional on-premises data warehousing solutions, which were designed to handle predictable, slow-moving, and predominantly structured data from internal transactional systems. However, these legacy systems tightly coupled storage and compute resources, requiring expensive hardware upgrades to scale and resulting in fragmented data silos. As organizations faced the influx of schema-less, semi-structured, and unstructured data from web applications, IoT devices, and social media, the rigidity of traditional Extract, Transform, Load (ETL) pipelines became a severe bottleneck. In response, the academic and industrial focus shifted toward cloud-native platforms, which democratized data analytics by removing the need for significant upfront infrastructure investments and offering elastic, on-demand resource allocation [3][14]. Platforms like the Snowflake Elastic Data Warehouse emerged during this transition to specifically address the shortcomings of on-premises architectures and early big data platforms, delivering a fully managed software-as-a-service (SaaS) solution optimized for cloud environments.



2. Research on Snowflake Architecture and Performance

A central theme in the literature regarding Snowflake is its innovative multi-cluster, shared-data architecture, which fundamentally decouples the storage layer from the compute layer. This architectural paradigm ingeniously blends the benefits of shared-disk and shared-nothing database models. Studies detail the platform's three-tier design: a centralized Data Storage layer (utilizing Amazon S3, Google Cloud Storage, or Azure Blob Storage), a Virtual Warehouses layer for isolated query execution, and a Cloud Services layer that handles metadata, security, and query optimization [4]. Performance evaluations frequently highlight Snowflake's state-of-the-art columnar, vectorized, and push-based execution engine. By utilizing a hybrid columnar storage format, Snowflake efficiently processes structured and semi-structured data (e.g., JSON, Avro, Parquet) natively, without requiring predefined schemas or complex data transformations. Furthermore, query execution speed is significantly enhanced by micro-partitioning and min-max based pruning, which allow the query optimizer to filter out irrelevant data chunks during execution, minimizing I/O overhead.

3. Scalability, Elasticity, and Cost Efficiency

The decoupled nature of Snowflake's architecture provides near-infinite, independent scalability of storage and compute resources, a feature widely lauded in comparative cloud studies. Research demonstrates that Snowflake's virtual warehouses can dynamically scale up to handle intensive workloads or scale out to support high concurrency without degrading performance or causing system downtime. Because these virtual warehouses operate autonomously, organizations can isolate diverse workloads—such as data engineering pipelines and ad-hoc analytics—preventing resource contention. Economically, Snowflake's pay-as-you-go pricing model charges users purely for the compute resources consumed per second, presenting a highly cost-efficient alternative to the continuous operational expenses of traditional hardware [9]. Nevertheless, recent studies indicate that without proper governance, uncontrolled query complexities and continuous high-intensity workloads can lead to unpredictable cost spikes. To mitigate this, researchers have explored applying predictive analytics and machine learning algorithms, such as XGBoost and time-series forecasting, to proactively anticipate resource usage, optimize virtual warehouse scaling, and refine query performance, ultimately yielding significant reductions in operational costs.

4. Snowflake in Business Intelligence and Real-Time Analytics

Business Intelligence (BI) is undergoing a transformative shift, transitioning from descriptive, retrospective reporting to proactive, real-time decision-making. The literature emphasizes Snowflake's role as the backbone for modern BI ecosystems, driven by its capacity to eliminate data silos and centralize vast datasets for real-time querying. A key enabler of this real-time paradigm is Snowpipe, a service that facilitates continuous, low-latency micro-batch data ingestion, allowing organizations to process streaming data from messaging platforms like Apache Kafka directly into analytic-ready models [10]. Additionally, Snowflake's Secure Data Sharing and Zero-Copy Cloning features revolutionize collaboration. Organizations can instantly create database clones for testing or share live data across departments and external business partners without moving or duplicating physical files, thereby drastically reducing latency and accelerating the time-to-insight.

5. AI Features: Snowpark ML, Cortex AI, and Predictive Analytics

The integration of Artificial Intelligence (AI) and Machine Learning (ML) directly into the data warehouse is one of the most prominent trends in recent literature. Historically, ML workflows required extracting large volumes of data from the warehouse into external computational environments, introducing latency, security risks, and governance complexities. Snowflake bridges this gap via Snowpark, a developer framework that allows data scientists to write complex feature engineering, data transformations, and model inference logic in Python, Java, or Scala directly within Snowflake's compute engine. By bringing the computational logic to the data, Snowpark streamlines the end-to-end ML pipeline, optimizing both execution speed and data security.



Furthermore, the advent of Snowflake Cortex AI has transformed the platform into a multimodal intelligence hub, embedding Large Language Models (LLMs) natively into the data environment [13]. Studies highlight the capabilities of Cortex Analyst, which translates natural language queries into precise SQL commands, democratizing data access for non-technical stakeholders. Additionally, Domain-AI agents leverage Scoped Retrieval-Augmented Generation (RAG) to provide context-aware, real-time insights based strictly on governed enterprise datasets. This targeted approach prevents the semantic hallucinations common in generic LLMs and provides accurate, verifiable analytics critical for highly regulated sectors such as procurement and compliance. Snowflake also maintains a thriving, open ecosystem, enabling seamless interoperability with external frameworks like TensorFlow, PyTorch, and Amazon SageMaker, empowering enterprises to deploy sophisticated predictive analytics and natural language processing models effortlessly.

6. Comparisons with Amazon Redshift, Google BigQuery, and Azure Synapse.

Comparative analyses in the literature frequently benchmark Snowflake against other leading cloud-native platforms, namely Amazon Redshift, Google BigQuery, and Azure Synapse. BigQuery is recognized for its fully serverless architecture and exceptional performance in real-time, ad-hoc analytics [6]. However, researchers note that BigQuery's pricing model can become expensive for organizations running highly complex or continuous ML workloads, whereas Snowflake's multi-cluster isolation allows for more controlled compute management. Amazon Redshift utilizes a Massively Parallel Processing (MPP) architecture that is highly cost-effective for extremely large-scale, structured data processing. Yet, Redshift struggles with the dynamic elasticity required for fluctuating workloads and is less efficient at natively handling semi-structured data compared to Snowflake. Azure Synapse offers deep integration within the Microsoft ecosystem, but Snowflake is consistently lauded for its cloud-agnostic deployment (available across AWS, GCP, and Azure), preventing vendor lock-in and supporting multi-cloud enterprise strategies. In contrast with legacy systems like SAP Business Warehouse (SAP BW), Snowflake demonstrates far superior elasticity and unstructured data handling, though SAP BW retains a stronghold in enterprises deeply entrenched in SAP ERP applications.

7. Enterprise Adoption and Industry Use Cases.

Snowflake's versatile architecture has spurred widespread adoption across various data-intensive industries. In the retail and e-commerce sectors, organizations leverage Snowflake to aggregate diverse data streams—such as web logs, CRM data, and transaction histories—enabling real-time customer segmentation, hyper-personalized marketing, and dynamic supply chain optimization. Within the financial services industry, the platform's low-latency processing and robust governance frameworks are critical for executing predictive models that detect fraudulent transactions, assess credit risks, and manage compliance reporting in real time.

The healthcare and bioinformatics domains present some of the most complex use cases for Snowflake [12]. Genomic sequencing, transcriptomics, and multi-omics analysis generate massive datasets, such as Variant Call Files (VCFs) that require specialized compression and rapid processing. Literature illustrates how researchers use Snowflake to integrate these datasets with public annotations (e.g., ClinVar, 1000 Genomes) for precise disease variant filtering and in silico drug discovery. Furthermore, pharmaceutical companies utilize Snowflake's Secure Data Sharing to create Data Clean Rooms, enabling secure, anonymized collaboration on clinical trial data across institutions while strictly adhering to HIPAA and GDPR regulations.

8. Key Findings, Common Trends, and Limitations in Previous Studies.

A synthesis of the reviewed literature reveals several overarching trends in modern data warehousing. The convergence of data lakes and data warehouses into unified "lakehouse" architectures is a prominent trajectory, driven by the need to manage raw, unstructured data alongside processed, structured data. Additionally, there is a clear paradigm shift toward AI-driven automation, where self-tuning databases and serverless architectures handle infrastructure scaling and performance optimization autonomously.



Despite these advancements, researchers identify several limitations and research gaps. While Snowflake's pay-as-you-go model is cost-efficient for dynamic usage, continuous, high-volume data storage and persistent ML training workloads can lead to escalating long-term costs compared to dedicated on-premises supercomputers. Vendor lock-in remains a peripheral concern; features like zero-copy cloning and proprietary data sharing protocols are highly vendor-specific, potentially complicating data portability and interoperability with external systems. Furthermore, as Generative AI and LLMs become deeply integrated into platforms like Snowflake Cortex, studies warn of challenges regarding AI governance, semantic gaps, and the necessity for explainable AI frameworks in regulated industries. Future studies are required to explore the optimization of complex multi-cloud deployments, the handling of high-velocity IoT data streams, and the refinement of adaptive security mechanisms that protect sensitive information without hindering analytical performance.

9. Security, Privacy, Compliance, and Data Governance

As organizations consolidate massive volumes of sensitive information in the cloud, rigorous data governance and security frameworks have become a focal point of recent academic and industrial research. Snowflake ensures end-to-end data security through a robust, four-level hierarchical key model (root, account, table, and file keys) rooted in hardware security modules like AWS CloudHSM. This cryptographic infrastructure automates key rotation and rekeying, providing transparent AES-256 encryption for data both at rest and in transit without requiring manual administrative intervention.

To address stringent regulatory compliance mandates such as HIPAA, GDPR, and SOX, studies highlight Snowflake's deployment of Role-Based Access Control (RBAC) integrated with Dynamic Data Masking [5][7]. This allows organizations to obscure Personally Identifiable Information (PII) contextually, ensuring that users only view data authorized for their specific roles. Furthermore, Snowflake's Secure Data Sharing capabilities have pioneered the implementation of Data Clean Rooms in the life sciences and financial sectors. These environments leverage differential privacy, allowing multiple institutions to collaborate on sensitive datasets—such as clinical trials—without physically duplicating or exposing the underlying raw data. Additionally, Time Travel features guarantee compliance and disaster recovery by preserving historical states of deleted or modified data for up to 90 days, enabling swift, granular data restoration.

10. Integration with BI, ETL, ELT, and Analytics Tools

The literature notes a significant paradigm shift facilitated by Snowflake's cloud-native architecture: the transition from traditional Extract-Transform-Load (ETL) pipelines to a more agile Extract-Load-Transform (ELT) methodology [3]. By natively ingesting raw data formats into VARIANT columns, data engineers can bypass rigid, upfront schema definitions and perform complex data transformations using the full computational power of Snowflake's parallel SQL execution engine.

Snowflake effectively bridges the gap between raw data storage and analytical consumption by maintaining seamless interoperability with a vast ecosystem of Business Intelligence (BI) and analytics tools, including Tableau, Power BI, and Apache Spark. Studies emphasize that the integration of automated data build tools like DBT, alongside real-time streaming services like Snowpipe, allows for the continuous, low-latency micro-batch ingestion of data. This unified integration strategy eradicates data silos, enabling BI dashboards to reflect near real-time operational insights rather than relying on stale, batch-processed reports.

11. Query Optimization, Concurrency, Caching, and Workload Management

Critical analyses of Snowflake's performance center on its decoupled, multi-cluster shared-data architecture and its custom execution engine. The execution engine is characterized as columnar, vectorized, and push-based, which avoids the materialization of intermediate results and maximizes CPU cache efficiency. Instead of relying on traditional,



overhead-heavy B-tree indices, Snowflake optimizes query execution through metadata-driven min-max pruning [4]. This technique automatically filters out irrelevant data micro-partitions during scans, drastically reducing I/O overhead. Workload management and concurrency are addressed through the elastic scaling of Virtual Warehouses, ensuring that resource-heavy tasks like bulk loading do not contend with complex analytical queries. To further minimize network latency, Snowflake utilizes a sophisticated multi-tiered caching architecture, combining local NVMe SSDs for caching frequently accessed file headers and columns, alongside a global result cache. Recent literature explores the integration of predictive analytics into workload management, demonstrating that machine learning models (such as XGBoost and ARIMA) can accurately forecast compute usage. By dynamically auto-scaling virtual warehouses based on these predictive models, enterprises have achieved substantial reductions in their overall compute expenditures without sacrificing query performance.

12. Semi-structured and Unstructured Data Analytics

A prominent differentiator for Snowflake in the literature is its native, optimized processing of diverse data types. Unlike legacy relational databases, Snowflake automatically infers the schema of semi-structured formats (e.g., JSON, Avro, Parquet, XML) and physically stores their most frequently accessed paths in a compressed, hybrid columnar format. This optimistic data conversion ensures that querying schema-less data is nearly as performant as querying strictly relational data, with benchmark tests showing minimal computational overhead. Beyond semi-structured data, recent studies emphasize Snowflake's capability to handle massive unstructured datasets crucial to fields like bioinformatics and healthcare [11]. The platform effectively stores and manages pathology-stained images, medical PDFs, and genomic Variant Call Files (VCFs) using built-in storage compression and external stages. Through SQL querying and integrated User-Defined Functions (UDFs), researchers can conduct high-performance filtering of millions of genetic variants directly within the data warehouse, demonstrating unparalleled flexibility in large-scale biological data analytics.

13. Future Trends: Generative AI and Autonomous Analytics

The convergence of cloud data warehousing and Artificial Intelligence marks the most dynamic frontier in current research. Future trends heavily feature the native embedding of Generative AI directly into the data warehouse ecosystem via Snowflake Cortex AI [13]. By integrating Large Language Models (LLMs) such as Llama 3.1 and Claude 4.5, Snowflake enables natural language interfaces (text-to-SQL) through Cortex Analyst, democratizing advanced analytics for non-technical stakeholders.

To mitigate the risks of algorithmic hallucination—especially in regulated sectors like finance and healthcare—the literature highlights the rise of Domain-AI and customized multi-modal agents. These scoped agents utilize Retrieval-Augmented Generation (RAG) to derive context-aware, verifiable insights strictly from governed enterprise data, rather than relying on generalized web knowledge. Looking forward, researchers predict a shift toward fully autonomous, self-tuning databases where federated learning and AI-driven automation continuously optimize query execution plans, resource scaling, and real-time anomaly detection, transforming the data warehouse into a proactive, intelligent decision-making engine.

III. CONCLUSION

In conclusion, this comprehensive review demonstrates that the Snowflake AI Data Warehouse has fundamentally redefined the enterprise data management landscape by overcoming the rigid limitations of traditional, on-premises architectures [3][8][14]. Its pioneering multi-cluster shared-data design, which completely decouples storage from compute, provides organizations with unprecedented elastic scalability, continuous availability, and cost-efficiency. Furthermore, Snowflake's profound integration with advanced analytics through features like Snowpark and Cortex AI has successfully bridged the gap between data storage and machine learning workflows. By enabling the execution of complex AI models directly within the data environment, Snowflake significantly reduces data movement latency,



enhances security, and empowers enterprises to transition from reactive reporting to proactive, real-time, and predictive decision-making.

Despite its transformative impact, several challenges remain in the adoption and optimisation of cloud-native data platforms. While the pay-as-you-go model is economically advantageous for fluctuating workloads, continuous and high-intensity machine learning operations can lead to unpredictable compute cost spikes if resources are not rigorously monitored and managed. Additionally, organizations face hurdles related to potential vendor lock-in from proprietary features (such as zero-copy cloning), complex data governance across decentralised teams, and the inherent risks of semantic gaps or hallucinations when deploying Large Language Models (LLMs).

Looking forward, the future scope of Snowflake and cloud data warehousing lies in the advancement of fully autonomous, AI-driven systems capable of self-tuning and dynamic resource allocation. Subsequent research and platform developments will likely focus on optimizing complex multi-cloud deployments, processing high-velocity data streams for Internet of Things (IoT) applications, and embedding sophisticated Domain-AI agents. Ultimately, as federated learning and explainable AI frameworks mature, Snowflake is poised to provide even more robust, compliant, and real-time intelligence ecosystems, ensuring secure and verifiable analytics for highly regulated sectors such as healthcare and finance.

REFERENCES

- [1]. M. S. Prasad, "A comparative study of Snowflake and SAP BW for data analytics," *International Journal of Technology, Management and Humanities*, vol. 11, no. 2, 2025, doi: 10.21590/ijtmh.11.02.07.
- [2]. J. K. Chinnathambi, "Amplifying big data utilization in healthcare analytics through cloud and Snowflake migration," *European Journal of Computer Science and Information Technology*, vol. 12, no. 6, pp. 15–23, Sep. 2024, doi: 10.37745/ejcsit.2013/vol12n61523.
- [3]. K. Allam, M. Ankam, and M. Nalmala, "Cloud data warehousing: How Snowflake is transforming big data management," *International Journal of Computer Engineering and Technology (IJCET)*, vol. 14, no. 3, pp. 156–162, Sep.–Dec. 2023.
- [4]. J. Wang and J. L. Kourik, "Data warehouse Snowflake design and performance considerations in business analytics," *Proceedings of ASBBS*, vol. 22, no. 1.
- [5]. R. Kashyap, "Data sharing, disaster management, and security capabilities of Snowflake: A cloud data warehouse," *International Journal of Computer Trends and Technology*, vol. 71, no. 2, pp. 78–86, Feb. 2023, doi: 10.14445/22312803/IJCTT-V71I2P112.
- [6]. B. P. R. Rella, "Data warehousing and its impact on machine learning model efficiency: Comparing Snowflake, BigQuery, and Redshift," *IRE Journals*, vol. 8, no. 10, Apr. 2025.
- [7]. D. P. Katsani, "Implementing robust data governance frameworks in the Snowflake platform," *International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT)*, vol. 11, no. 1, pp. 274–281, Jan.–Feb. 2025, doi: 10.32628/CSEIT25111226.
- [8]. G. Nookala, "Integration of AI and machine learning with Snowflake: Enabling advanced analytics and AI-driven insights in cloud environments," *International Journal of Artificial Intelligence, Data Science, and Machine Learning (IJAIDSML)*, vol. 6, no. 1, pp. 147–155, Jan. 2025, doi: 10.63282/3050-9262.IJAIDSML-V6I1P116.
- [9]. S. H. Kukkuhalli, "Optimizing Snowflake enterprise data platform cost through predictive analytics and query performance optimization," *International Journal on Science and Technology (IJSAT)*, vol. 15, no. 4, Oct.–Dec. 2024.
- [10]. W. K. Akram, "Revolutionizing business intelligence through cloud computing, AI/ML, and Snowflake database optimization," Dec. 2020.



- [11]. P. Borra, “Snowflake: A comprehensive review of a modern data warehousing platform,” *International Journal of Computer Science and Information Technology Research (IJCSITR)*, vol. 3, no. 1, pp. 11–16, Jan.–Dec. 2022, doi: 10.5281/zenodo.11617628.
- [12]. T. Koreeda, H. Honda, and J.-i. Onami, “Snowflake data warehouse for large-scale and diverse biological data management and analysis,” *Genes*, vol. 16, no. 1, p. 34, Dec. 2024, doi: 10.3390/genes16010034.
- [13]. V. S. K. Indurthy, “Snowflake and domain-AI real-time intelligence for enterprise data warehousing,” *International Journal of Science, Research and Technology (IJSRAT)*, vol. 9, no. 2, Mar.–Apr. 2026, doi: 10.15662/IJSRAT.2026.0902007.
- [14]. B. Dageville et al., “The Snowflake elastic data warehouse,” in *Proc. 2016 Int. Conf. on Management of Data (SIGMOD '16)*, San Francisco, CA, USA, Jun. 2016, doi: 10.1145/2882903.2903741

