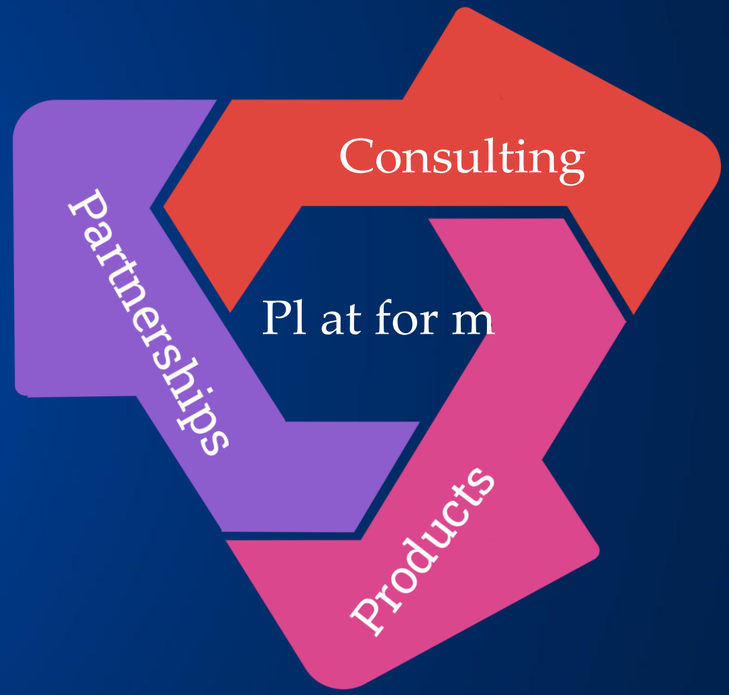
AulNova Techsoft

Projects Delivered to Valued Clients
Our Trusted Partnerships & Successful Projects
A Showcase of Our Work
Exceeding Expectations: for Our Esteemed Clients

AI first company focussed to deliver data driven customer value



Customer Value As North Star

Prioritizing customer satisfaction above all, guiding business decisions towards delivering meaningful value.

Cultivating a Culture of Innovation

Foster an environment where creativity is encouraged, risks are embraced to drive transformative change

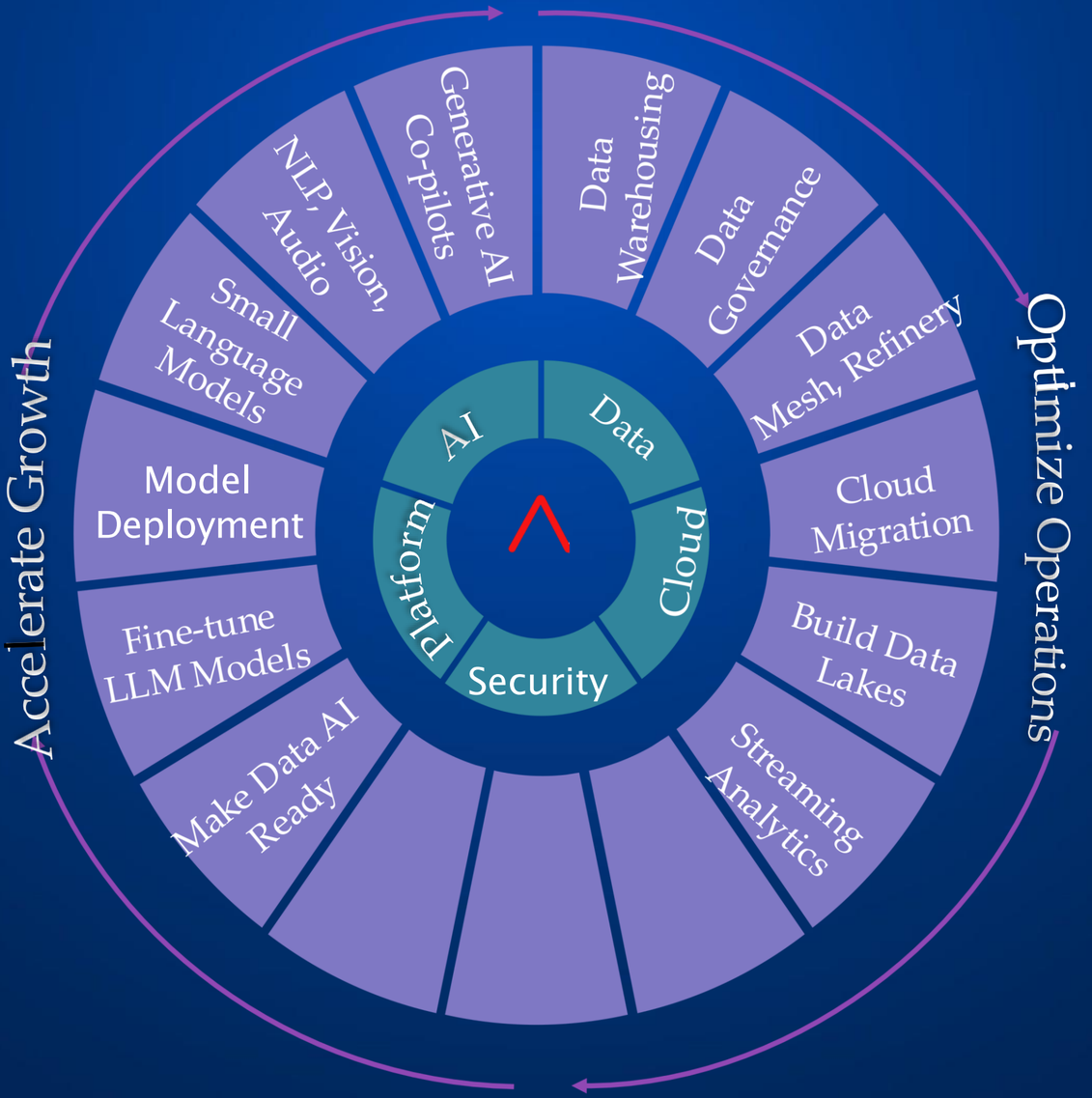
Hybrid Intelligence

Focus on rigorous capability building to integrate business expertise and technological depth

Team Well Being Over Organisation Grandneur

We value individual contributions, prioritize team satisfaction, and achieve collective success through

Our Services



Warehouse Management

6 warehouses for different workloads (ETL, BI, Data Science, ML, ad hoc reporting)

Query Optimization

Long-running queries taking over 600 seconds

Data Storage Efficiency

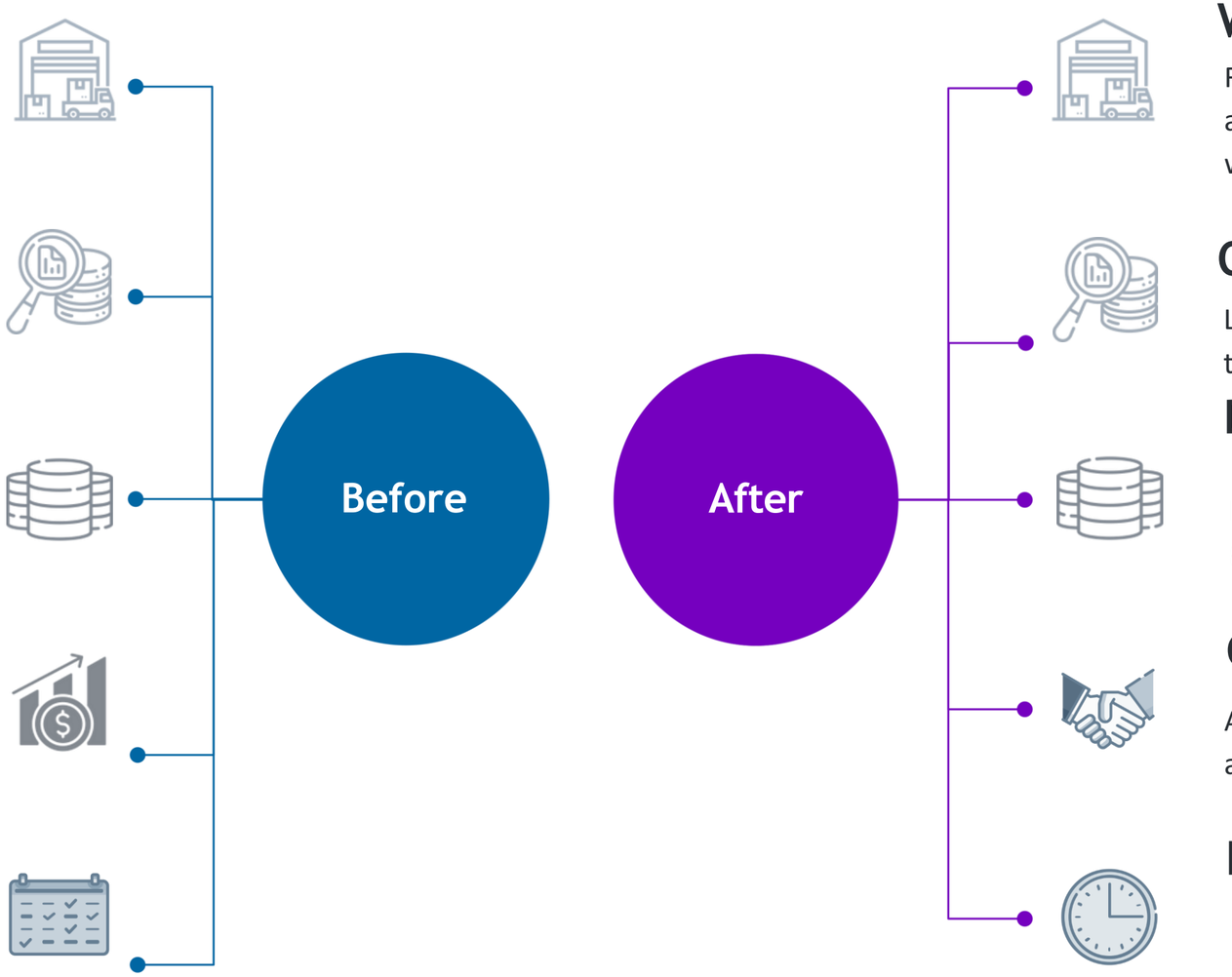
Uncompressed data storage format inflating costs

Cost Reduction

Substantial increase in Snowflake bill month-on-month post-migration

ETL & BI Performance

ETL loads taking 48-55 hours to complete



Warehouse Management

Reduced to 4 warehouses with accurate scaling for diverse wo rk lo a ds

Query Optimization

Long-running queries reduced to 300-400 seconds

Data Storage Efficiency

Implemented columnar compressed formats (e.g., Parquet) to reduce storage costs

Cost Reduction

Achieved a total annual savings of approximately $385,000

ETL & BI Performance

ETL loads taking 24-28 hours to complete

Finance

Objective: Design and Deploy a RAG chatbot to query and retrieve answers across PDFs, excel files, and audio recordings

Approach

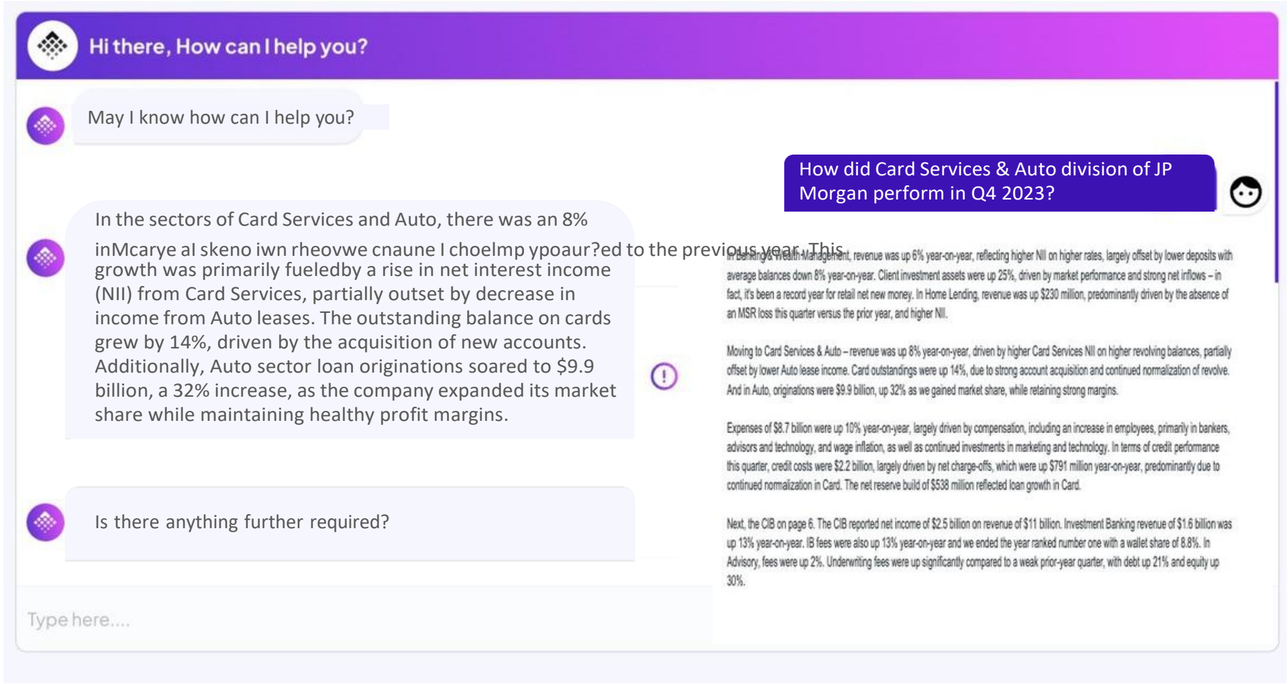
Create a custom vector DB dynamically to ensure role-based access to information for different users

Transcribe audio files to text documents, embed them and add to the vector DB based on request

Imp act

50%

Reduction in time, as reported by Investment Bankers, to identify and query through data



Supply Chain

Objective: Maximize cash flow through advancedinventorymanagement and demand prediction

Prob l em

Excessive inventory levels tied up capital, while stockouts resulted in lost sales o pp o rtu n itie s

Impact Delivered

15%

Reduction in capital tied up in inventory by optimizing inventory turnover rate, decreased inventory holding cost

Approach

Data Collection: Gathered historical data, customer demographics, seasonal trends, promotional activities, weather and economic indicators

Integrated the upstream workflow activities (supplier details, lead times) to determine optimal reorder quantity and reorder point for each product

Created a model to evaluate customer creditworthiness and mitigate the risk of late payments and bad debts

Integration layer is established to combine outputs of both models into unified framework and capture human feedback

Pharma

Objective: Extractmedicalentitiesfrom a BILLION pharmadocuments

Challenges with LLM Models:

- Skyrocketing Inference Cost
- High Latency Time
- Data Privacy& Security
- High Resource U ti l izat io n

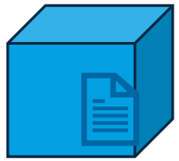
Results Delivered:

	Solaris	GPT-4
Model Type	SLM	LLM
Parameters	355M	1.7 Trillion
No of Rows	1000	1000
Inference Time	17 seconds	40 minutes
Inference Cost	$0.005	$0.45

Ap proa ch

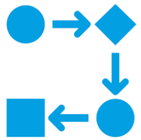
Time for customization: 2 Weeks

Annotatedthe pharma docume nts using Annotation Studio



Fine-tuned our proprietary SLM on the annotated datasets

BuiltBiomarker co-pilot on top of the fine-tunedmodel



Integrated the Copilot in the existingworkflow with no data movement

Frugal Innovation

Faster Time to Value

Operationalize AI

Health care

Objective: Predict the probability of developing chronic diseases in next 30 days

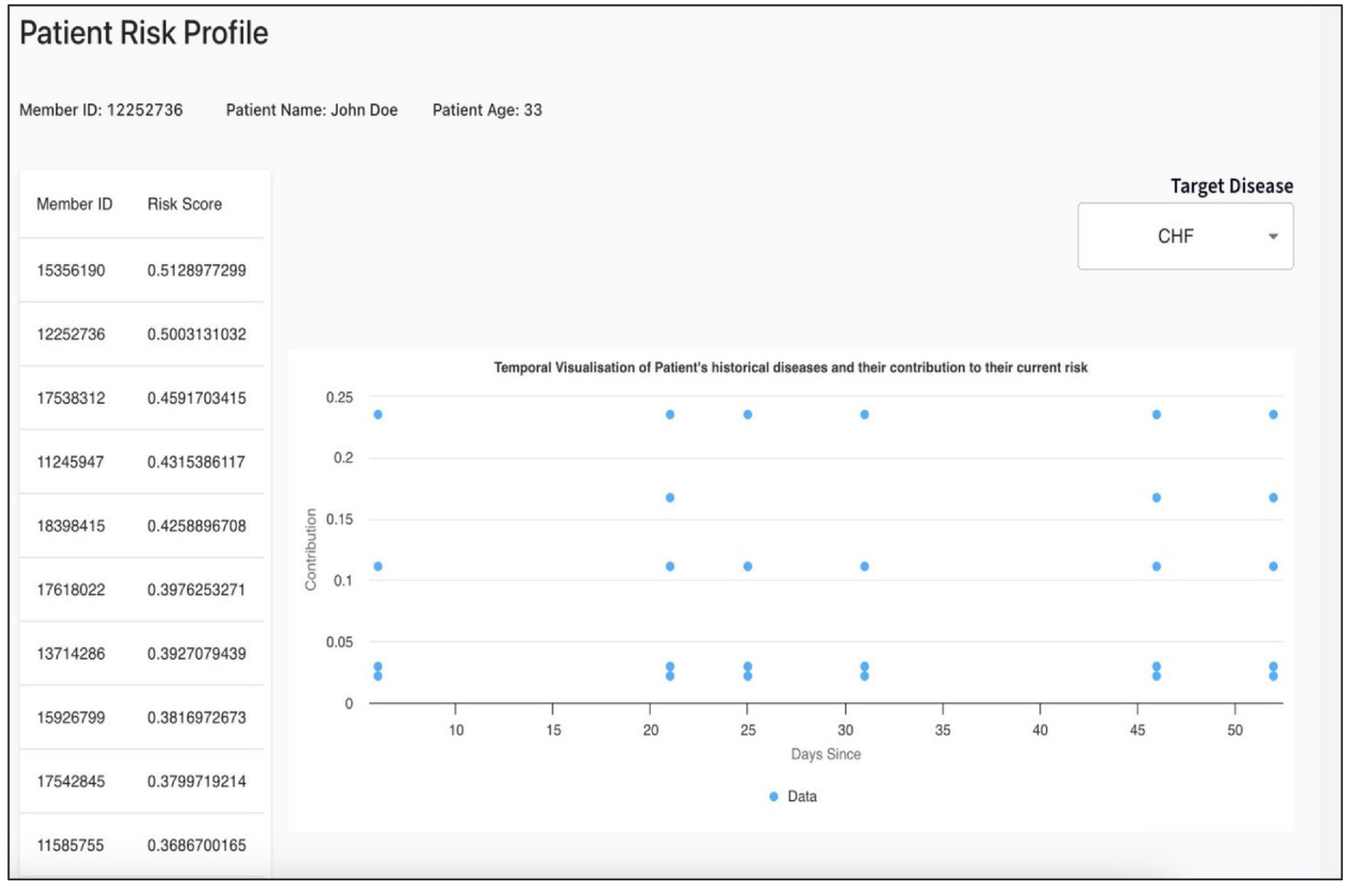
Imp act

15%

Lift in detecting chronic condition compared to traditional models

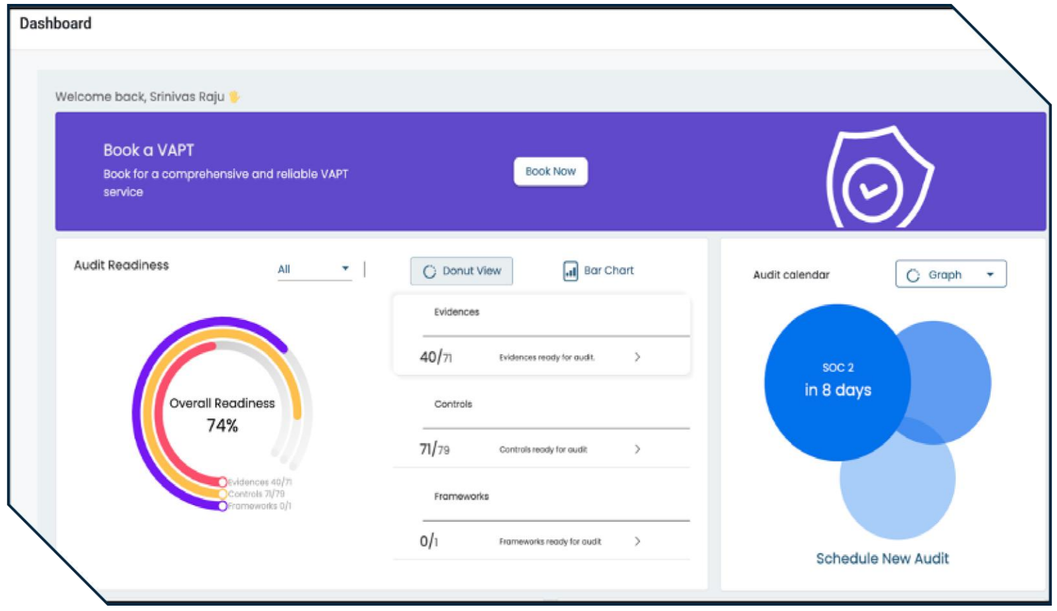
70%

Reduction in costs for operationalizing AI

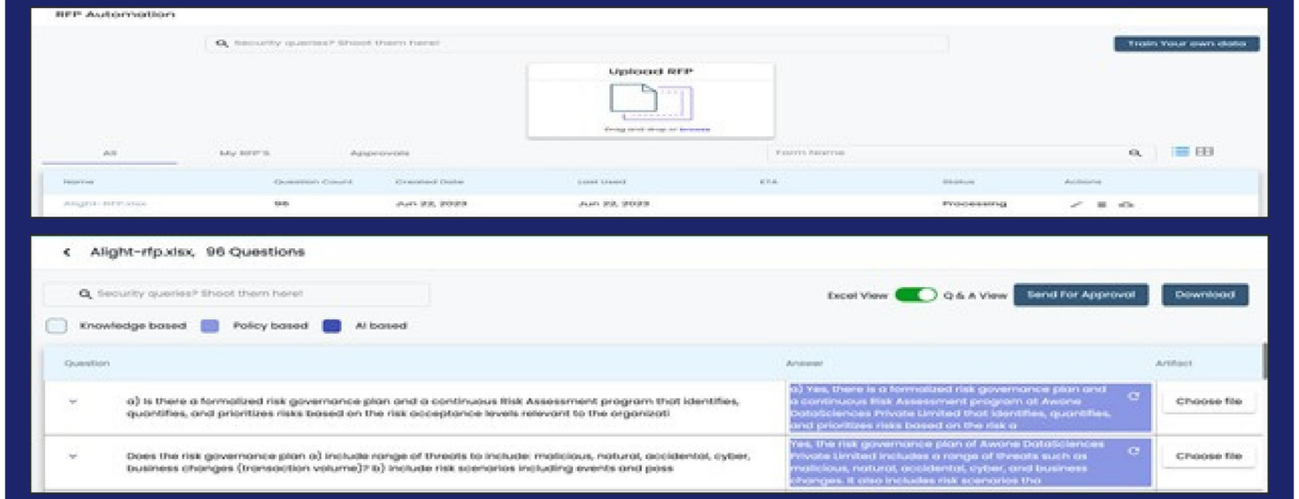


Compliance

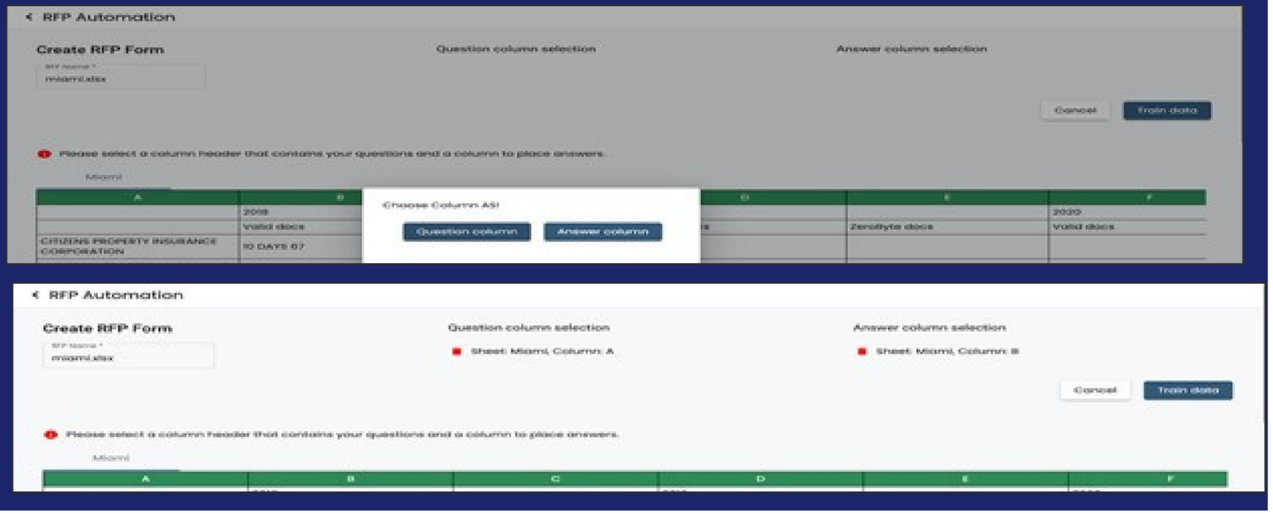
Objective:A Compliance automation and Audit management platform, tasked to automate the process of responding to security RFPs.



Deployment



Fine Tuning



Finance

Objective: Improve the operational efficiency of the Brokers with real-time data

Imp act

Improved Efficiency: 35% reduction in manual research time through automated data collection and integr ation

Volatility Index: Improved the accuracy of volatility index based on broader market sentiment and trends

Real-time update: Through continuous data integration, near real-time data is made available for quick decisions

Approach

Data Aggregation: Collection of vast datasets from financial news outlet, market DB, social m ed ia

Natural Language Processing: Utilizing NLP to convert unstructured to structured data and capture sentiment in real-time

Testing & Optimization: Rigorous backtesting with historical data to refine the model capability and ensure re li ab il ity

Integration: Seamlessly integrating the model with existing brokerage platform for real-time analysis

Oil & Gas

Objective: Developan oil& refinedproducts demand forecastingmodel

Challenges with LLM Models

$10M estimateonlyfor fine-tuning the LLM model on the historicaldata

Hallucination: LLM models startedhallucinating afterfine-tuning on a large dataset

InferenceCost:Due to the large sizeof the LLMs, it provedto becostlyto keepthe model in production continously

Fully-private: Due to sensitivityof data, model has to befullyowned, controlledand governed

Approach

Data Annotation: Annotedthe documentsusingour Label studio and createda data registryof all datasets

Model Fine-tuning: Builta task-specificSmall Language Model over the organizationspecific data

ReinforcementLearning with Human Feedback: Implementeddata feedback loop to capturehuman feedback to improvemodel performance

Hallucination:Implementedguardrailsto preventany occurenceof potential hallucinations

InferenceCost: Inferencecostis projectedto reduceto 1/10thcomparedto LLM, throughauto-scalingof GPUs& distributed computing

Data Services

Migration to Cloud

Simplified database migration from on premise to any public cloud

Minimal downtime migration and highly-performant at scale systems

Highly available , stable and secure that customers trust for their mission-critical workloads.

Building DataLakes

Collect all types of structured and unstructured data

Store raw data

Elimination of data silos

Democratized access to information via a unique, centralized view of data across the org anization

Data Warehouse as a Service

Performant, highly available SQL data warehouse like Snowfla ke

Accelerating adoption and implementation

Multitude of data connectors

Secure - Complaint role-based access control on data

Streaming Analytics

Built streaming pipeline using serverless architectures and auto-scala ble

Run mission critical workload with sub second latency

Power real-time analytics with artificial intelligence

Case study

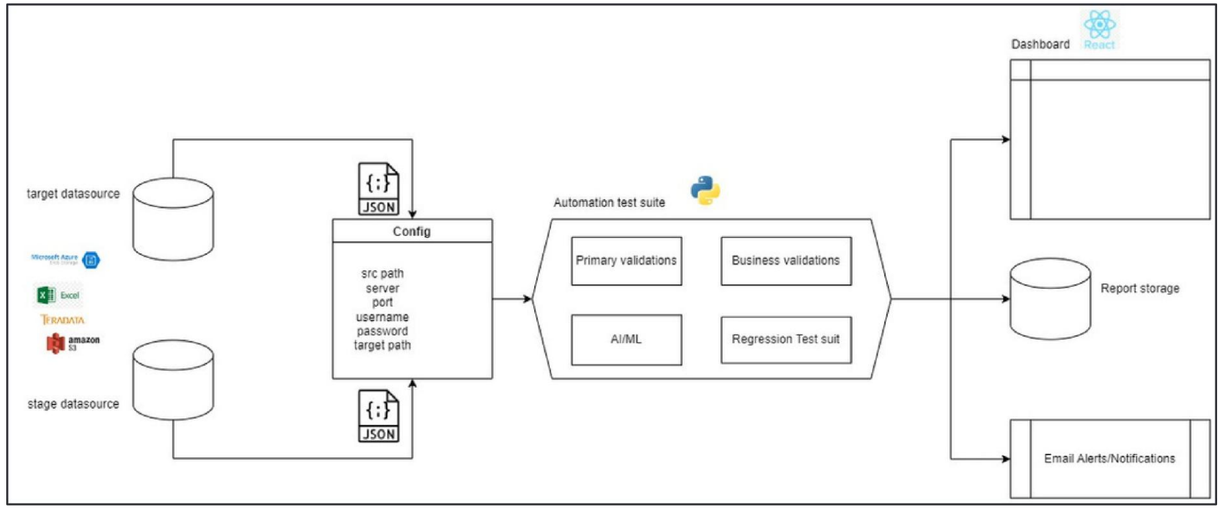
Data Quality

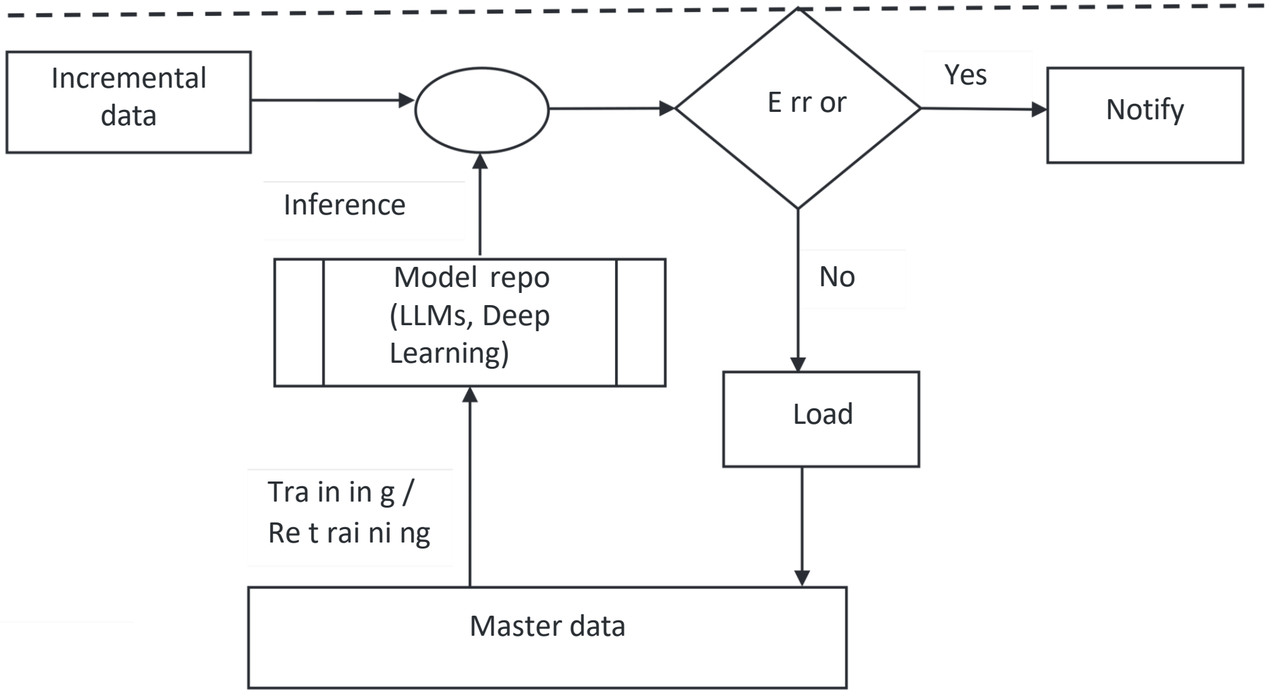
Data Validation Framework

Generative AI based framework that generates test cases, SQLs based on historical test plans, mapping documents.

These test cases cover

- Primary Validations
- Business Validations
- Regression Test suite
- And AI validations





Anomaly Detection Framework

Data Enrichment:

Address standardization, Missing value imputation

Anomaly Detection

Gen AI based techniques to Identify anomalies in string variables
Identify duplicates

Case study

Data Services

Australian Football League

Centralized data platform:

30% reduction in man-hours; streamlined data collection, storage, and management from disparate sources

360-degree view of player history:

Track player performance beyond AFL league, laying the foundation for future analytics and informed decision-making

Northcott

Data process flow visibility:

Availability of data to downstream from integration layer enabled the teams to track the cost at project level

Cost reduction:

Transparency in the process helped the total cost of hiring to joining across organization and teams

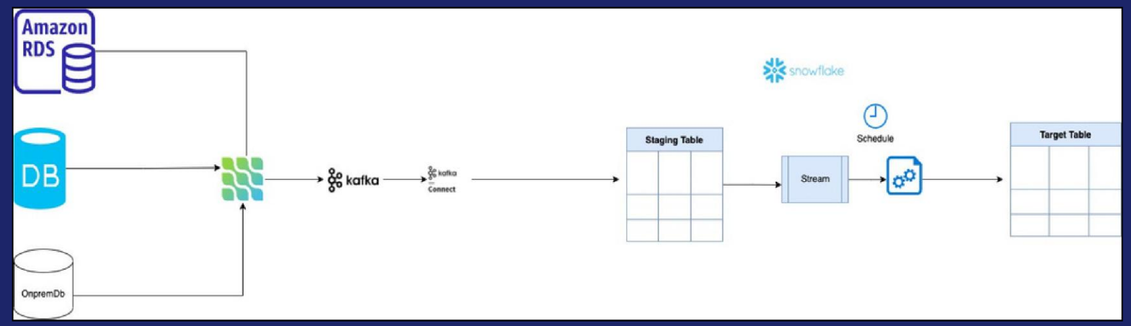
Snowflake Accelerator

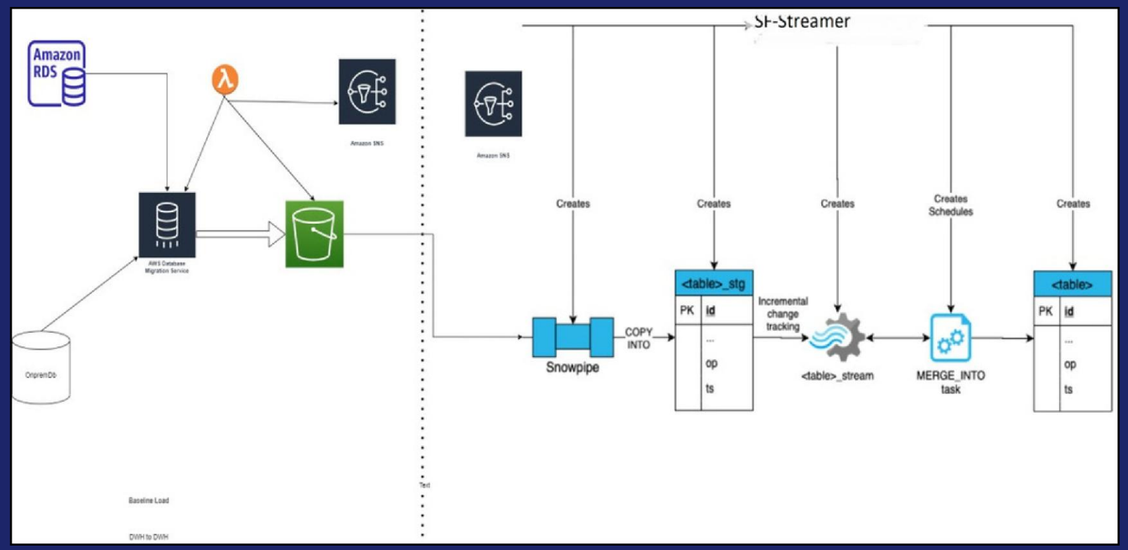
Empowers enterprise to continuously process and ingest data onto snowflake database with just the click of a button.

- User interface and APIs for snowflake Ingestion along with Data Integrity reports
- Uses Debezium/DMS to transfer the source data comprising of databases, object storage and streams
- Compatible to source data from on-perm and any other cloud provider with continuous ingestion
- Set up infra using automated DevOps pipeline, ingest data from Kafka to snowflake
- End to End automated process for data migration with out any worry on infra, balancer, API and security

Success story

- Time to delivery has significantly improved for one of our c li en ts.
- Building real time analytics has become faster.





HealthCare

Objective: Optimizing snowflakeworkloadand warehouseperformance, improvisingcostand TCO factors

Challenge

A Fortune company recently migrated from on-premises Teradata and Netezza systems to Snowflake. The migration involved over 4,000 data pipelines, sourcing from 15+ systems, with a warehouse size of approximately 80 TB (compressed). However, post-migration, Snowflake costs exceeded initial estimates by more than 40%.

Workload Segmentation: Utilized multiple warehouse clusters without optimizing for specific workload types

Data File Formats: Migrated data without compressing files, leading to higher storage costs.

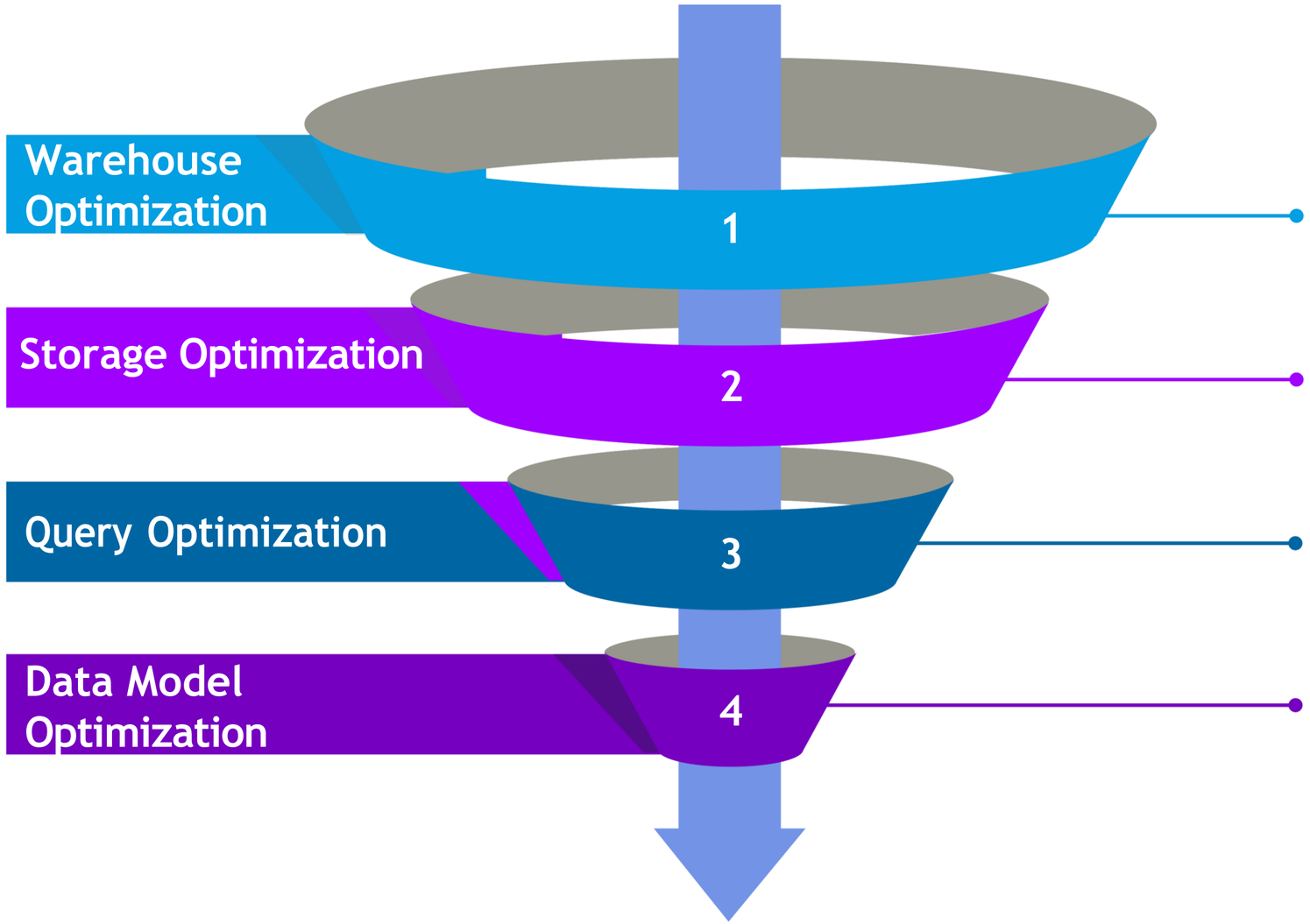
Inaccurate Sizing: Inaccurate workload estimation resulted in improper cluster sizing, escalating expenses.

Inefficient Object Layer: The object layer wasn't implemented with best practices, causing unnecessary overhead.

Data Modeling Flaws: The migration suffered from an inefficient data model with poor distribution and clustering strategy.

Lack of Optimization: No focus on implementing Snowflake-specific optimization techniques, leading to suboptimal performance and cost inefficiency.

Top-to-Bottom Approach



Optimize query performance by **minimizing queue depth** and memory issues. **Right-size clusters**, and enhance retrieval and **concurrency management**

Use compressed formats and clustering to **enhance data access efficiency**. Implement **search optimizations** for faster point-lookups in datasets.

Improve **partition pruning**, scan efficiency, and data loading processes. Use automation tools like **Airflow, monitor concurrency.**

Fine-tune cluster keys and optimize large table designs. Enhance reporting with aggregated models and speed up access with **multi-column clustering**

Scenarios

Scenario 1:

Problem: Several large tables (1 TB+) had time travel enabled and were created with clustering keys. Frequent re-clustering operations led to increased storage consumption due to the accumulation of data in the fail-safe feature.
Solution: Implemented enhanced query performance monitoring to assess the necessity of re-clustering. Unnecessary re-clustering was pruned, and automatic clustering was suspended where it wasn't required. This resulted in significant cost savings by optimizing the use of clustering operations.

Scenario 2:

An analysis of workload history revealed that more than 40% of system resources on warehouses were consistently under-utilized. Vertical scaling was the default choice for queries involving large volumes of data, yet only 5% of queries were long-running, and these were mostly ad-hoc.

Solution: Implement workload-based dynamic scaling by adjusting warehouse sizes according to the nature of the queries. For frequent, short-running queries, smaller warehouses were used, while vertical scaling was reserved for the infrequent, long-running, high-volume queries. Additionally, implemented auto-suspend/resume policies to ensure warehouses are only active when needed, preventing idle time and maximizing resource efficiency. This optimization reduced compute costs while maintaining performance for high-volume queries.

Scenario 3:

Inefficient query execution patterns were leading to high compute costs and long processing times. Several factors contributed to this, including large data volumes being processed unnecessarily, excessive use of sorts, complex self-joins, and the reliance on views that slowed down query performance. Additionally, there was a lack of optimization around query pruning, caching, and throughput, resulting in under-utilized Snowflake features and increased resource consumption.

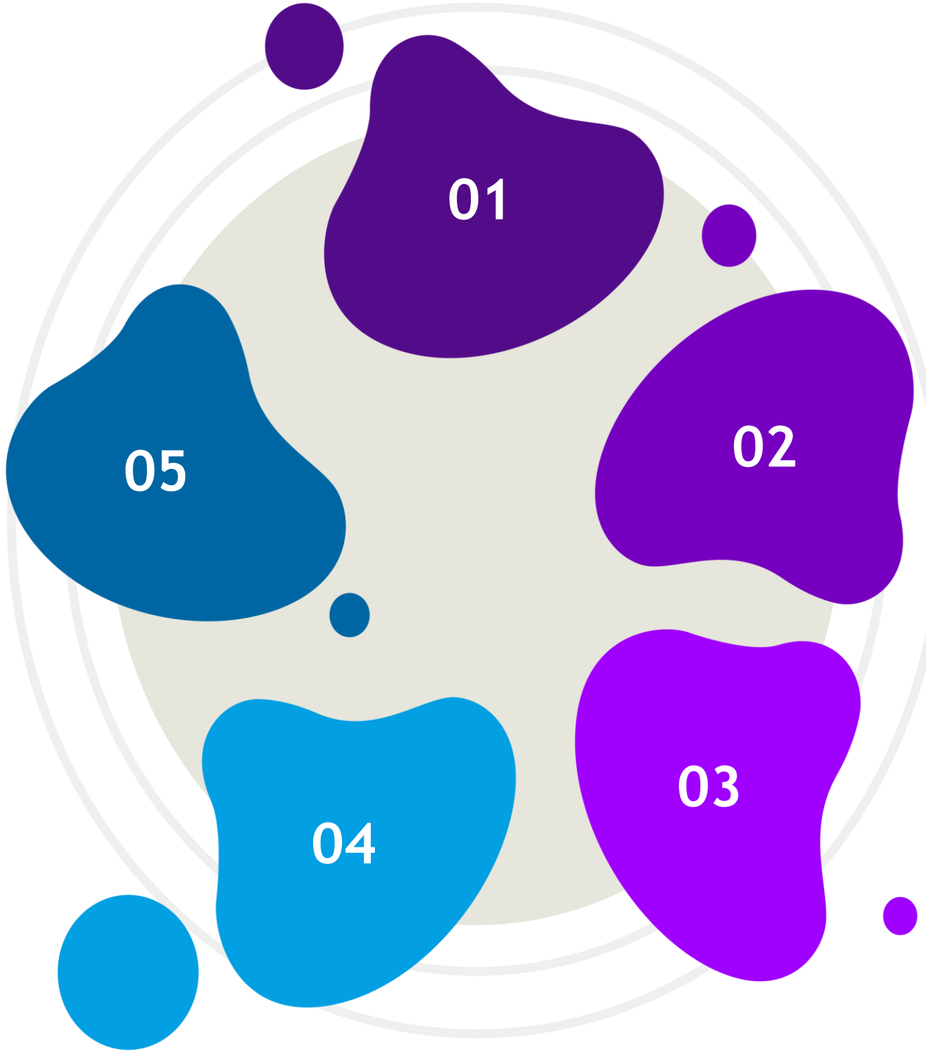
Solution: Optimized queries by implementing data pruning, clustering join columns, and using pre-aggregated tables to reduce data processed. Eliminated unnecessary sorts, replaced self-joins with window functions, and simplified views to improve performance. Ensured efficient use of query caching and managed throughput to optimize resource usage, significantly reducing compute costs and improving query speed.

Large-scale Implementation

Deploy production-ready Snowflake warehouses at scales of 60 TB to 250 TB

Expert Team

Leverage certified Snowflake Solutions Architects and Data Engineers for implementation



Snowflake Migration

Seamlessly migrate from on-prem systems like Teradata and Netezza

Catalog Migration

Use Snowflake Accelerator for efficient and seamless catalog migration

Optimization & Tuning

Optimize and fine-tune Snowflake for peak performance

i

https:\\aulnovatechsoft.com

mailto:contact@aulnovatechsoft.com