



Snowflake Builders Blog: ...

Transforming Business Search with AI: How Markaaz and Snowflake's Cortex Search Enhance Speed, Quality, and Innovation

Rustin Scott · [Follow](#)

Published in Snowflake Builders Blog: Data Engineers, App Developers, AI/ML, & Data Science · 5 min read · Feb 25, 2025



30



Markaaz, a leading global provider of business verification, monitoring, and data enrichment solutions, is partnering with Snowflake to bring state-of-the-art AI search experiences to our customers. Leveraging Cortex Search, Snowflake's native search solution for vector and lexical search, we are bringing the power of semantic search to our customers, improving their search experience through increased match rates.

Business search is at the core of what Markaaz does. Our customers leverage search on our platform to perform activities like business verification and monitoring. Given that, search quality and search speed are among the most

important metrics we track, and we are always looking for ways to improve our performance to provide a better experience for our global customers.

In 2024, we made the decision to power search over Markaaz's Business Directory of over 540 million global business records with Cortex Search. This decision, which was coupled with the migration of our entire data platform to Snowflake, enables us to provide our customers across the globe with high search quality out of the box, with little tuning required. With minimal effort, we've increased recall on real customer queries, thereby reducing the amount of time our customers require to perform business verification activities.

Our decision to leverage Cortex Search was all about the following:

- **Simplifying the architecture:** powering search with Snowflake, our unified data platform, means we don't need to procure a 3rd-party solution and egress data to it
- **Increasing quality and performance:** Cortex Search gave us a +12% boost in recall so our customers can get better matches on their queries, faster. Our innovative matching algorithm is now even higher quality than before.
- **Searchable data on arrival:** Business data changes quickly, and Cortex Search allows us to provide searchable data to our customers almost immediately without the need for cumbersome reindexing of the entire directory.
- **Unlocking innovation and AI at scale:** Cortex Search gives us great quality and performance out of the box. We can provide our customers with a delightful search experience without staffing up a large team of

search and infrastructure experts. Using Snowflake we are also able to focus our energy on innovation and customers.

Simplifying the architecture

Originally, search at Markaaz was powered by a 3rd-party enterprise search solution managed opaquely by a team of consultants. As we moved to Snowflake as our data warehousing and engineering platform, we saw an opportunity to consolidate solutions in our data stack with Cortex Search.

Today, the Markaaz Business Directory contains over 540M business records across the globe, each of which contains up to 200 data elements. Because this data is stored in and governed with Snowflake, we saw Cortex Search as a natural product to power our search experience.

We were astonished to find we could get up and running with Cortex Search on our entire 540M+ row directory with a single SQL query. In a matter of hours, we had built a search experience superior to our current solution in quality with no extra infra provisioning, ETL pipeline setup, or quality experimentation required.

Not only that, but **Cortex Search enables us to reduce the complexity in our data stack in two key ways:**

- **Remove an extra third-party search service from our architecture** — no need to manage another ETL pipeline and worry about secure data movement. Cortex Search is fully managed in Snowflake, along with the rest of our data.
- **Remove requirement for an extra OLTP database** — no need to maintain and query a separate OLTP database that we were previously using to

enrich search results with metadata (up to 200+ fields) per query. Cortex Search allows us to store all search data and metadata in a single Service, so it's now one database / query instead of two.

Increasing quality and performance

For our customers, search quality is everything. We are the solution that can verify business that others can't, and our success in this domain is driven by our innovative matching algorithms and broad business directory. We take search quality seriously because it directly impacts our customers' revenues by allowing them to acquire more customers while minimizing risk.

Our prior search solution provided lexical-only search. Cortex Search is a “hybrid” search service which leverages both vector (semantic) search and traditional keyword search, along with a semantic reranking step, to provide state-of-the-art quality out of the box. Cortex Search, powered by Snowflake's arctic-embed vector embedding model, allowed us to easily take advantage of recent advancements in the GenAI-based search domain without managing any additional AI infrastructure.

To assess the quality of Cortex Search, we compiled a “golden” dataset of pairs of real customer queries and desired business match results. We then evaluated our existing solution (the baseline) against Cortex Search without any tuning as well as Cortex Search with a novel feature called soft query boosting for weighting important fields.

In short, we saw a **5% increase in hitrate@5 over baseline without any tuning**, and a **7% increase in hitrate@5 over baseline with soft query boosting**.

Not only this, but we saw Cortex Search bring end-to-end search times down from ~3.5 seconds to <500ms. This reduction has a critical impact on our user experience, reducing wait times for both our API and Business Lookup experiences. This means our customers can increase efficiency in their business onboarding processes and free up their compliance teams to focus on higher-value tasks.

Searchable Data on Arrival

The Cortex Search platform that makes data searchable within minutes of arrival. This cutting-edge solution enables organizations to query and access the most up-to-date information with unprecedented speed, addressing the growing demand for immediate data insights in today's fast-paced business environment.

By leveraging real-time insights, businesses can stay ahead of onboarding and compliance requirements, optimizing operations. The platform's ability to handle continuous data streams while maintaining high availability and low latency positions it as a game-changer for industries ranging from finance and insurance to e-commerce and compliance.

Unlocking innovation and AI at scale

As a small, scrappy team of innovators, we are always looking for ways to increase efficiencies in our data stack so that we can spend more time on our main objective — obsessing over customer needs.

With Cortex Search, we've significantly improved the quality and performance of our prior solution without having to hire a team of AI/ML and infrastructure engineers to tune and operate the solution. The managed

infrastructure, simple user interface, and high quality out of the box means that Cortex Search is doing the job of multiple engineers. This allows us to treat search as a mostly solved problem and redirect our team's focus on higher-value problems like analytics and new products for customers. This re-focusing on product innovation and leveraging AI at scale is a key motivating factor behind our decision to move search onto Snowflake.

In summary, with Cortex Search we integrated a much stronger search solution into our data platform, got faster search, improved search quality, and greater transparency into a crucial part of our product.

Snowflake

Performance

Search

Data Modernization



Published in Snowflake Builders Blog: Data Engineers, App Developers, AI/ML, & Data Science

Follow

8.4K Followers · Last published 17 hours ago

Best practices, tips & tricks from Snowflake experts and community



Written by Rustin Scott

Follow

1 Follower · 1 Following

Grizzled tech wizard, once slaying network dragons. Now, this 25-year IT veteran dives into data lakes, wielding SQL spells and Python charms.

No responses yet





Elliott Botwick

What are your thoughts?

More from Rustin Scott and Snowflake Builders Blog: Data Engineers, App Developers, AI/ML, & Data Science



 In Snowflake Builders Blog: Data Engine... by Jee...

Snowflake integrates with Amazon S3 Tables

At Snowflake, we're committed to help organizations accelerate their open...

Mar 14  33  2



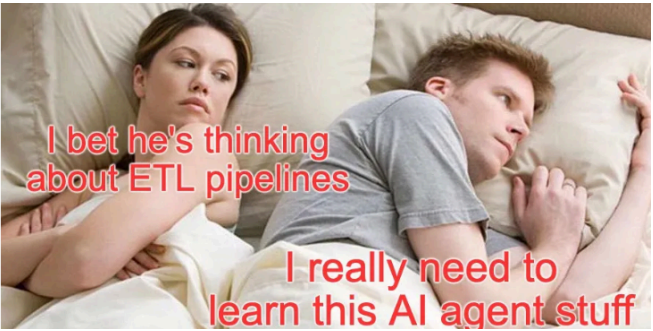
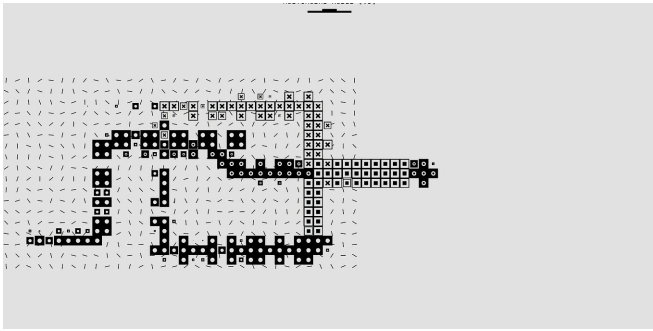
 In Snowflake Builders Blog: Data En... by August...

Snowflake Security—Service Accounts

Learn about how designate User Accounts as Services Accounts for an Improved Security...

Aug 23, 2024  65





 In Snowflake Builders Blog: Data Engin... by Haebi...

Bringing Gen AI to ETL: Redefining Data Engineering through LLMs

The Future of ETL is Here (through Snowflake)

Oct 22, 2024  152  3  

 In Snowflake Builders Blog: Data Engin... by Vino ...

Agentic AI: A Buzzword or a real deal? Why should you care?

What makes an AI agent? Why data engineers should learn about orchestrating AI agents?

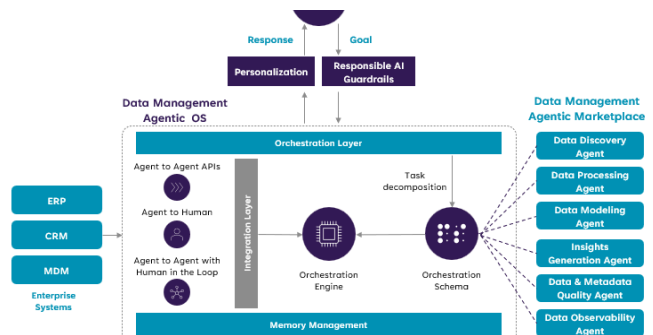
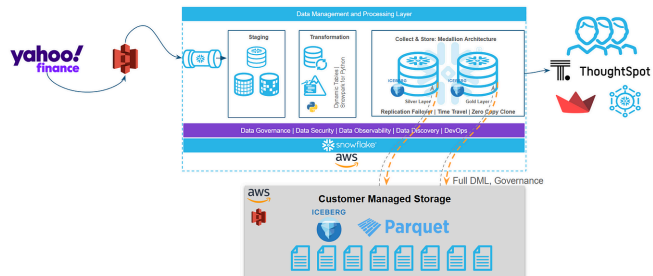
Feb 27  40  1  


See all from Rustin Scott

See all from Snowflake Builders Blog: Data Engineers, App Developers, AI/ML, & Data Science

Recommended from Medium

Iceberg Lakehouse Implementation Using Snowflake

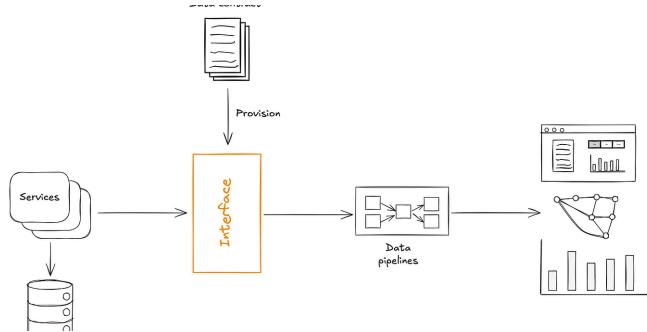


 Divyansh Saxena

[2025] Building a Data Lakehouse Using Snowflake Iceberg —...

Ever wonder how iceberg tables work with snowflake? How can architecture be built...


4d ago  4  3  

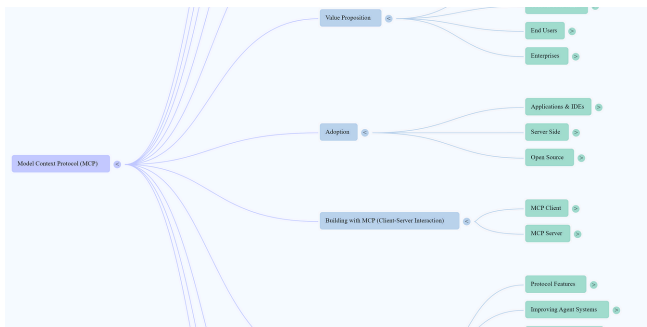


 Andrew Jones 

Integration without coupling, with data contracts

Data contracts has always been about solving the problem of tightly coupled data...

Feb 20  3  



 Dr. Nimrita Koul 

The Model Context Protocol (MCP) —A Complete Tutorial

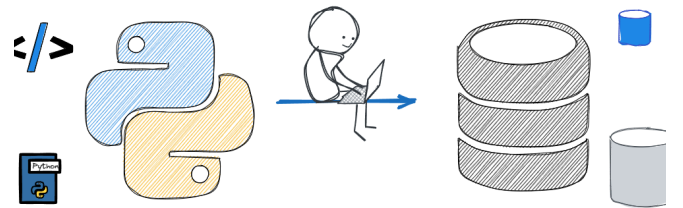
 In AI Advances by Debmalya Biswas


Agentic AI for Data Engineering

Reimagining Enterprise Data Management leveraging AI Agents

Mar 23  625  16  


Python Meets SQLGlot: Seamlessly Parse and Convert SQL Queries

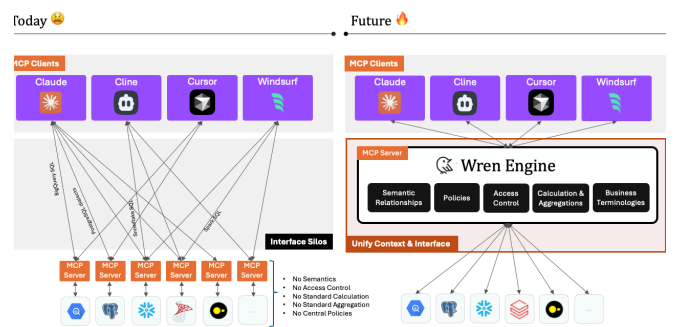



 Abhishek Das

Python Meets SQLGlot: Seamlessly Parse and Convert SQL Queries

Have you ever found yourself jumping between SQL dialects—MySQL, SparkSQL,...

3d ago  6  



 In Wren AI by Howard Chi

Fueling the Next Wave of AI Agents: Building the Foundation for Futur...

Anthropic released the Model Context Protocol(MCP) in Nov. 2024.

6d ago  255  1



How Wren Engine Powers the Semantic Layer for AI-Driven Workflows—Unlocking Context...

3d ago  134



See more recommendations