Deloitte. | snowflake

Trustworthy Agentic AI at Scale with Data Agents

AI and data solutions from Deloitte and Snowflake

Deloitte has been a long-time proponent and developer of **trusted AI foundations** for our clients. The advent of agentic AI, which runs mostly autonomously, makes trust essential. Agentic AI leverages large language model (LLM) outputs and automates desired actions without constant human intervention. This creates opportunity for transformational change but can introduce greater risk at a faster pace.

To deploy trusted agentic AI at scale, your data management practices must be solid.

Data agents can help.

Snowflake has been adding features to support the development of agents natively within the platform, maintaining a protected environment where AI tools and data reside together behind the firewall. Snowflake users can now accelerate agentic AI at scale with support from Cortex Agents—a new data agent that provides seamless interaction between models and data while maintaining accuracy, trust, and compliance.

Accuracy with speed
Deliver high accuracy out-of-the-box and meet quality requirements

Trust and security
Comply with security and governance policies in real time

Governed data access
Access structured and unstructured data sources across systems

Built on top of enhancements to Cortex Analyst and Cortex Search, teams using Snowflake can create Cortex Agents to streamline agentic application data access and orchestration for more reliable AI-driven decisions. (Future Cortex Agents will be even easier to deploy, empowering business analysts.)

With Cortex Agents, Deloitte and Snowflake can help accelerate trustworthy, enterprise-wide agentic AI solutions that deliver:

Enhanced efficiency: Automate routine tasks such as data entry, cleansing, and validation, freeing teams to focus on strategic activities that drive growth

Scalability: Effectively manage large-scale data environments, adapting to increasing volumes and complexities

Improved data quality: Continuously monitor and validate data sources, detecting anomalies, correcting errors, and updating records in real-time

Real-time insights: Continuously analyze data for real-time insights, and identify trends that enable swift responses to market changes, customer preferences, and emerging opportunities

Improved accuracy and precision: Combine the flexibility of LLMs with the precision of traditional programming to make more informed decisions based on context and real-time data

Trustworthy Agentic AI

Traditional data management methods are increasingly insufficient given exponential data growth. Many enterprises face overwhelming data sources, from structured databases to unstructured social media feeds, and manual processes can be time-consuming and error prone. Agentic AI automates these processes, helping ensure data integrity and offering real-time insights. Leveraging advanced machine learning and natural language processing, these intelligent agents can efficiently manage and analyze vast data amounts. The integration of Snowflake's AI Data Cloud and launch of Cortex Agents, along with Deloitte's experience and Trustworthy AI™ framework, can optimize these processes for efficiency and innovation.

SNOWFLAKE OBSERVABILITY FEATURES

Model behavior can change over time due to input drift, stale training assumptions, and data pipeline issues. Snowflake AI Observability for LLM Apps* supports the Deloitte Trustworthy AI approach by enhancing evaluation and trust in LLM applications during development. With AI Observability, you can better meet compliance requirements and assess relevance, groundedness, and bias, alongside traditional quality metrics such as latency. Keep RAG systems accurate with the RAG triad, a feature set devoted to evaluating context relevance, groundedness, and answer relevance to avoid hallucinations and other potential violations of trust for RAGs. Once in production, Snowflake ML Observability allows you to track the quality of models across multiple dimensions, such as performance, drift, and volume.

DELOITTE TRUSTWORTHY AI

Deloitte's Trustworthy AI framework provides a methodical way forward for training and operating AI solutions you can trust. This includes engineering solutions that address:

Bias
Perform regular bias audits and implement moderation and filtering to avoid unequal or unfair recommendations.

Hallucination
Continuously monitor outputs and performance improvements. Execute additional fine-tuning and prompt engineering.

Privacy and security
Continuously enforce infosec leading practice policies for what, how, and when access is provided to external models.

Inappropriate behavior
Put robust QA processes in place to proactively minimize risk of inappropriate responses and maintain customer trust.

Emergent abilities
Monitor models for abilities or novel behaviors that can appear with scale and time and implement fine-tuning.

Cost and accountability
Leverage FinOps capabilities to account and budget for the costs of hosting and fine-tuning large models.

Deloitte and Snowflake Can Power Your Agentic AI Ambitions

Jason Eichenholz
Snowflake Lead Alliance Partner
Deloitte Consulting LLP
jeichenholz@deloitte.com

Rupesh Dandekar
Snowflake Chief Technology Officer
Deloitte Consulting LLP
rudandekar@deloitte.com

Sara Hennessey
Alliances Vice President
Deloitte Consulting LLP
shennessey@deloitte.com

See our latest blog for insights.

*As of March 14, 2025, AI Observability for LLM Apps was in private preview. Check the Snowflake website for updates.