

SNOWFLAKE LOGO SNOWFLAKE CORTEX AISQL: YOUR AI-POWERED SQL SUPERPOWER! 🚀

Cortex AISQL (AI_SQL) transforms Snowflake into an AI powerhouse, enabling you to harness Generative AI directly within SQL. It makes interacting with, analyzing, and transforming unstructured or multimodal data (text, images, audio, documents) seamless and secure, all without leaving your Snowflake environment. If you know SQL, you unlock AI!







🚀 WHAT IT IS: AI AS SQL

- **SQL-Native GenAI:** Embeds LLMs directly into SQL functions.
- **In-Snowflake Processing:** Runs AI within your account, against your data, using existing security and governance.
- **SQL First:** If you know SQL, you can use AI_SQL – no new languages or complex setups needed.
- **Multimodal Data:** Works with text, images, documents, audio, and structured tables.
- **Managed LLMs:** Leverages leading models (OpenAI, Anthropic, Meta, Mistral, DeepSeek, Snowflake Arctic) hosted within Snowflake Cortex.
- **"Talk to Your Data":** Enables natural language querying and AI-powered insights.
- **AI Operators:** Integrates into data engineering workflows for transformation.






User Interaction Flow: Visualize how AI_SQL translates natural language input into actionable SQL queries and delivers results.

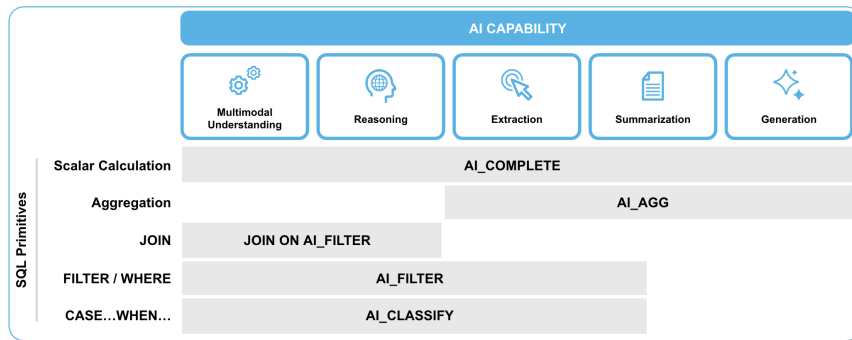
CORE CAPABILITIES & IDEAS

-  Summarize, classify, extract data, and detect sentiment/patterns.
-  Parse documents (PDFs, invoices) & transcribe audio/video.
-  Generate narratives, explanations, and deeper insights.
-  Create embeddings for semantic search, similarity, and clustering.
-  Leverages familiar SQL syntax for AI tasks.
-  Integrates with data engineering workflows as "AI Operators".
- **Design Goal:** Empower users to perform AI tasks without separate infrastructure.
- **Key Idea:** If you know SQL, you can use AI_SQL.

AI_SQL FUNCTION FAMILIES

Over 16+ functions, categorized for clarity, usable in SQL with full RBAC and Snowflake governance:

Category 	Key Functions 	Use Cases 
Text Generation	AI_COMPLETE, TRY_COMPLETE, PROMPT	Generate summaries, explanations, replies, transform text. Supports advanced options. `TRY_COMPLETE` adds robustness. `PROMPT` enables template reuse.
Classification & Sentiment	AI_CLASSIFY, AI_FILTER, AI_SENTIMENT	Categorize text, detect tone, filter risk, classify content by natural language.
Aggregation & Summaries	AI_AGG, AI_SUMMARIZE_AGG	Summarize text across rows, handling large volumes beyond context windows.
Embeddings & Similarity	AI_EMBED, AI_SIMILARITY	Vectorize text for semantic search, clustering, and identifying similar data points.
Document & Audio Processing	AI_PARSE_DOCUMENT, AI_TRANSCRIBE	Extract content from PDFs, invoices (OCR), and transcribe audio/video.
Extraction & Translation	AI_EXTRACT, AI_TRANSLATE	Pull entities (names, dates, amounts) and translate multilingual data.
Utilities	AI_COUNT_TOKENS, TO_FILE	Estimate token costs and prepare staged files for AI functions.



AI Capabilities Mapped: Visualize how AI capabilities align with SQL primitives and AI functions.

🔒 SECURITY & GOVERNANCE: YOUR DATA STAYS PUT

- **Data Residency:** Data never leaves Snowflake.
- **No External Dependencies:** No need for API keys or external hosting.
- **Access Control:** Via SNOWFLAKE . CORTEX_USER role and RBAC.
- **Cost Management:** Token-based metering (input+output), tracked via usage views. Use AI_COUNT_TOKENS for planning.
- **Seamless Integration:** Works with Tasks, Streams, Snowpark for automation.
- **Managed LLMs:** Uses leading models hosted within Snowflake.
- **Data Privacy:** User data is not used to train public models.
- **Governance:** Fully governed under your Snowflake account.

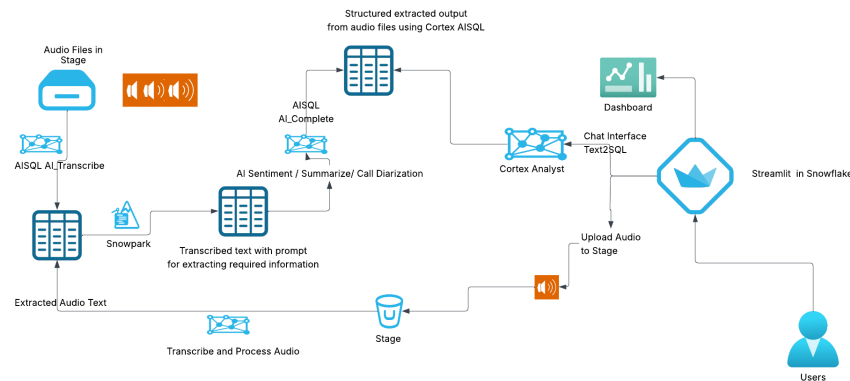
☀️ TYPICAL USE CASES

- **Customer Experience:** Classify/summarize feedback, sentiment analysis.
- **Call Analytics:** Transcribe calls (AI_TRANSCRIBE), extract info, detect sentiment.
- **Document Processing:** Parse PDFs/invoices (AI_PARSE_DOCUMENT), extract details (AI_EXTRACT).
- **Social Media:** Translate (AI_TRANSLATE), filter content (AI_FILTER), cluster posts via embeddings.
- **Conversational BI:** Ask natural language questions (e.g., "top pain points?") for AI insights.
- **Data Engineering:** Use as AI operators in ELT pipelines for enrichment.
- **AML/KYC:** Extract entities from documents, classify risk cases.
- **Content Analytics:** Summarize articles, extract keywords.
- **Localization:** Translate and adapt multilingual content.
- **Unstructured Joins:** Integrate text/image columns for similarity joins.

- **Feature Generation:** Create embeddings for SQL logic/models.

Workflow showing audio files processed through AI_TRANSCRIBE, Snowpark, and AI_COMPLETE for call diarization and sentiment analysis.

Audio Processing Workflow: Visualize an end-to-end audio processing workflow using AISQL.



Detailed Audio Processing Workflow: Illustrates a comprehensive data pipeline for audio files using various AISQL functions and Snowflake components.

💡 EXAMPLE: CHAINED AI FUNCTIONS FOR SMART AUTOMATION

Combine functions for sophisticated workflows, e.g., summarize then analyze sentiment:

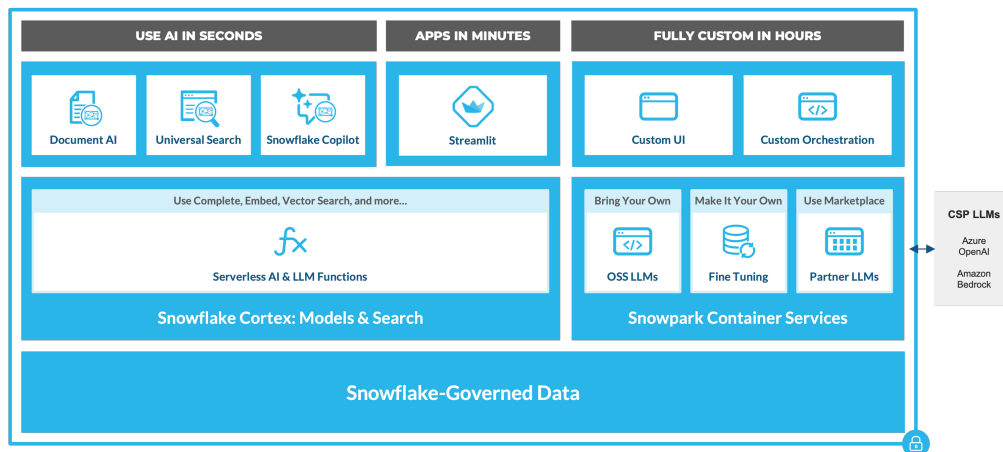
```
SELECT
  region,
  AI_SENTIMENT(AI_SUMMARIZE_AGG(feedback_text)) AS sentiment_summary
FROM swiggy_feedback
GROUP BY region;
```

This efficiently summarizes and analyzes feedback in one SQL query.

🧩 HOW IT FITS WITH SNOWFLAKE ML & AI

- **Cortex AI_SQL:** GenAI for unstructured data via SQL.
- **Cortex Agents/Analyst/etc.:** Higher-level AI experiences built on AISQL.
- **ML Functions:** Predictive analytics on structured data.
- **Snowpark ML:** Custom model training.

Together, they form Snowflake's complete AI Data Cloud. ☁️



Snowflake AI/ML Architecture: Illustrating the layered approach from governed data to integrated AI capabilities.

⚡ KEY BENEFITS

- **Democratized Analytics:** Natural language querying for business users. 🌐
- **Reduced Manual Effort:** Automates data prep and insight generation. ⌚
- **Seamless Integration:** Embeds AI into ELT and reporting workflows. 🔗
- **Security & Compliance:** Data remains within Snowflake's governed environment. 🔒
- **Scalability:** Leverages Snowflake's elastic compute for AI workloads. ⚡
- **Cost Efficiency:** Token-based metering and optimized LLM usage. 💰
- **Unified Platform:** Combines structured (ML Functions) and unstructured (Cortex AISQL) data intelligence.
- **Developer Productivity:** Simplifies AI implementation via SQL.
- **Performance Gains:** 30–70% improvements reported on certain workloads.
- **Conversational Analytics:** Makes analytics conversational.

🚩 ONE-LINE DEFINITION (INTERVIEW-READY)

Snowflake AI_SQL runs Generative AI in SQL—querying, summarizing, classifying, and extracting insights from unstructured data inside Snowflake, securely powered by Cortex LLMs.

⚠️ LIMITATIONS, CAVEATS, AND WHAT TO WATCH

- **Latency:** AISQL calls are heavier than plain SQL; use judiciously.
- **Cost:** Monitor token usage; optimize prompts and use aggregations/caching.
- **Hallucinations:** AI can produce incorrect outputs; treat as assistive and validate critical results.

- **Preview vs. GA:** Check docs for function/region availability.
- **Prompt Sensitivity:** Standardize and version prompts for production.

DESIGN PATTERNS AND BEST PRACTICES

- **Prompt Engineering:** Be explicit, use few-shot examples, and leverage structured extraction instructions (JSON).
- **Chaining Functions:** Create pipelines (e.g., Parse → Extract → Embed).
Patterns: Summarize → Sentiment; Translate → Classify.
- **Robustness & Safety:** Use TRY_* variants, guardrails, and always validate critical outputs.
- **Data Engineering Integration:** Use within Streams/Tasks for scheduled transforms. Call from Snowpark/apps. Combine AISQL with ML Functions for hybrid pipelines.

This comprehensive guide consolidates all provided details into concise bullet points with key visuals and descriptions. Please let me know if there's anything else you need!