

|||| ClickHouse

# meetup

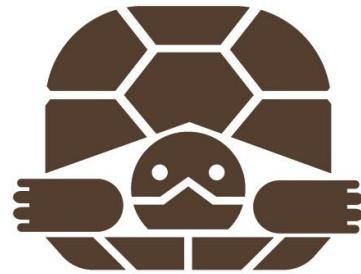
**TAO BIN office**

Bangkok, Thailand

September 16, 2025 at 6:30 PM ICT



# Thank you to our host!



TAO BIN

• BEVERAGE •



# **Maruthi Prithivirajan**

## **Solution Architect @ ClickHouse**

**Say hi!**



**Read my blogs!**



# Agenda

**01**

Intro to ClickHouse

**02**

Uniqueness of ClickHouse

**03**

AI/ML with ClickHouse

# ClickHouse

A Quick Overview

Feb 2025

||||· ClickHouse

# ClickHouse

2009  
*Prototype*

2012  
*Production*

2016  
*Open Source*

2021  
*ClickHouse Inc.*

2022  
*ClickHouse Cloud*

**The Most Popular Analytics Database on the Planet**

#1  
*Analytics DB on DB-Engines*

Over  
**40,000**  
*GitHub Stars*

Over  
**200,000**  
*Community Members*

# What is ClickHouse ?

ClickHouse is an **Open-Source**, columnar **OLAP** database  
Designed for **Blazing fast** analytics of massive volumes of data

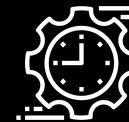
**1**

**Speaks SQL fluently**



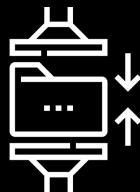
**2**

**Processes data very fast**



**3**

**Highly efficient storage**



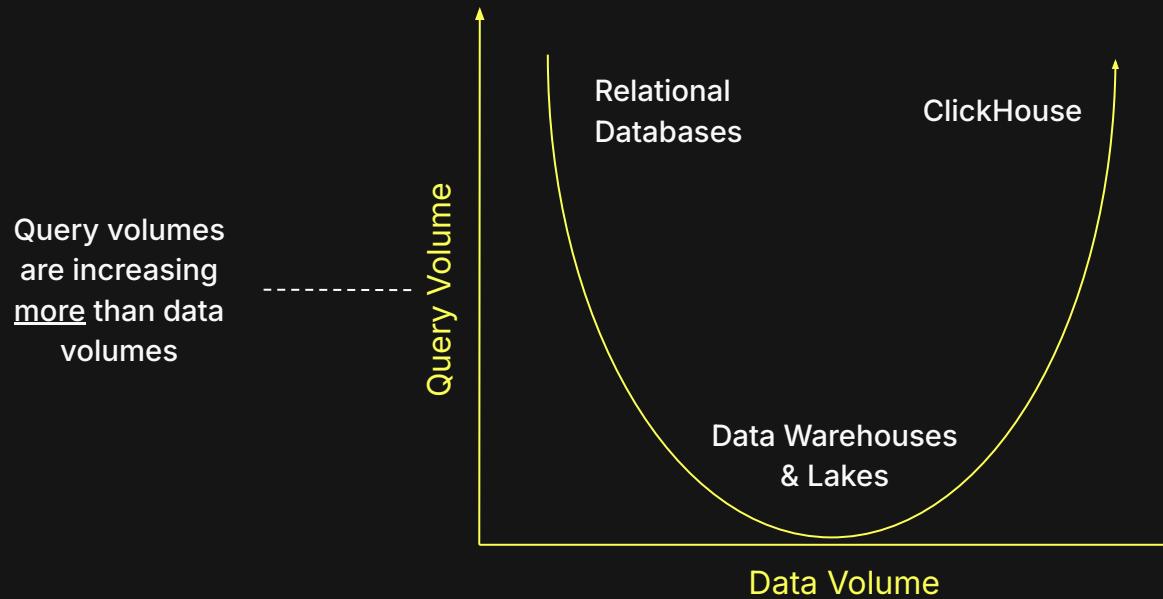
**4**

**Easily scalable to any size**



# Build for the future

Skating to where the puck is going



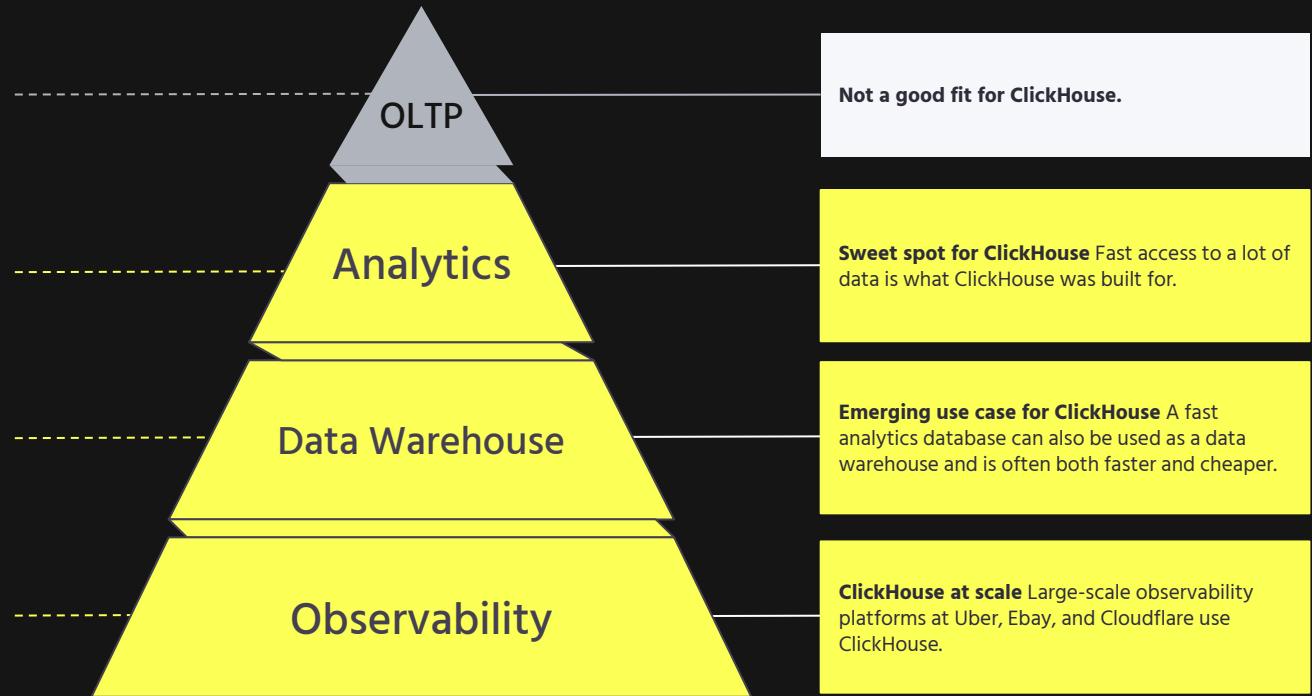
# Hierarchy of Data

This is where your most important data lives. Customer names and addresses, payment info, open orders, etc.

Customer facing and internal analytics: Your real-time interactive dashboards and APIs on top of customer transactions and activity, order history, etc.

Where all business data goes, anything that is useful today or maybe will be at some point in the future: All orders, transactions, random datasets, etc.

The “digital exhaust” of all your applications and infrastructure: Logs, metrics, traces. Used for real-time monitoring and debugging.



# Who uses ClickHouse?



Bloomberg



coinbase



SONY

ROKT



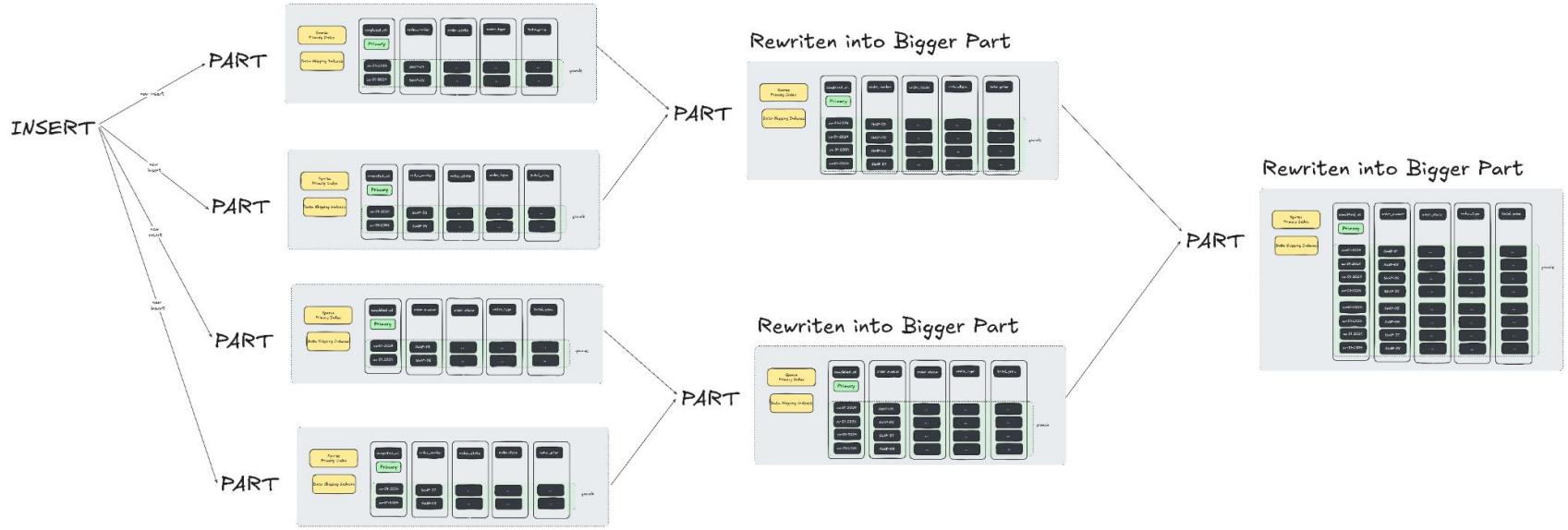
zomato

... & thousands more



# What Makes ClickHouse Unique?

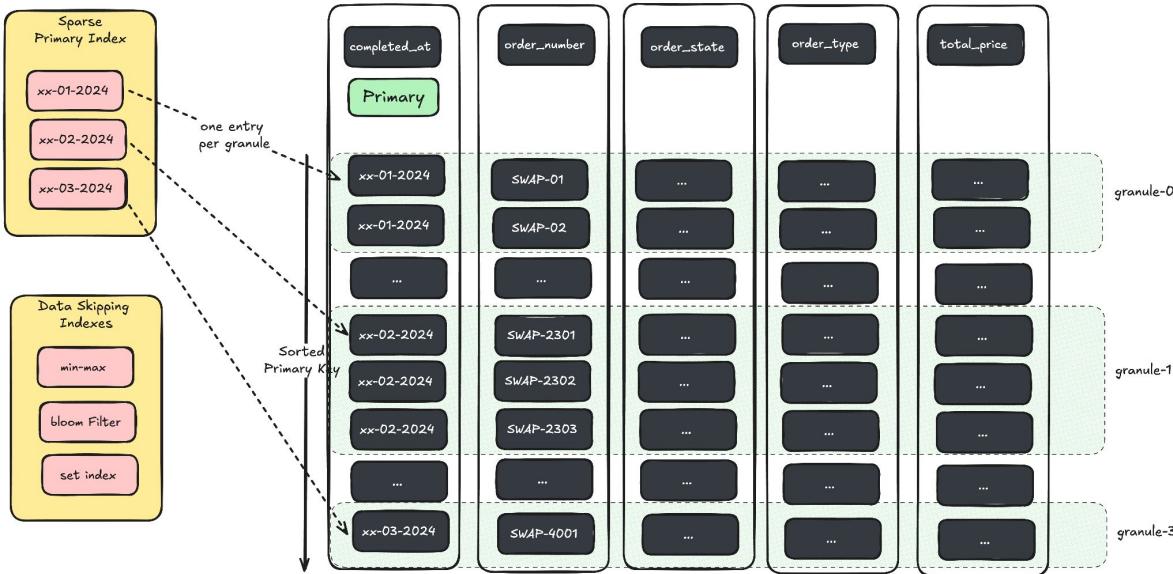
# Why: Efficient Merging



**Parts:** Each time data is inserted, a new part is created. Multiple parts can accumulate over time, which can slow down query performance.

To improve performance and reduce the number of parts, **ClickHouse** periodically runs merging operations.

# Why: Multiple Layers of Indexes



- **Granules** are chunks of rows, and **ClickHouse** groups rows into granules based on the index\_granularity setting.
- **Sparse Primary Index:** **ClickHouse** uses the sparse primary index to quickly jump to the relevant granules, skipping over large portions of the data that don't need to be read.
- **Skip Indexes:** These allow **ClickHouse** to skip entire granules if it can determine from the index that no rows in the granule satisfy the query's filter.

# Use Cases

## Real-Time Analytics

- Applications, Dashboards, APIs
- Customer facing
- Interactive
- Querying ClickHouse directly

## Data Lake & Warehouse

- Business data for internal use
- Use any BI software on top (Tableau, Power BI, Superset, Metabase, Amazon QuickSight)
- Move between ClickHouse and Iceberg, Parquet, Delta Lake, etc.

## Observability

- Logs, metrics, traces
- Built-in integration with OpenTelemetry
- Grafana as a UI
- Unify on SQL as the query language

## Machine Learning & GenAI

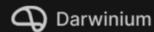
- Feature Store
- Vector Search
- LLM Observability
- MCP Server

# Use cases



## Logs, events, traces

Monitor with confidence your logs, events, and traces. Detect anomalies, fraud, network or infrastructure issues, and more.



**zomato**

**ebay**

**resmo**

**runreveal**



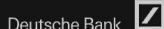
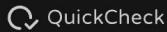
## Real-time Analytics

Power interactive applications and dashboards that analyze and aggregate large amounts of data on the fly. Run complex internal analytics in ms, not mins or hrs.



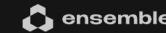
## Business intelligence

Interactively slice and dice your data for analysis, reporting, and building internal applications. Evaluate user behaviors, ad and media perf, market dynamics, and more.

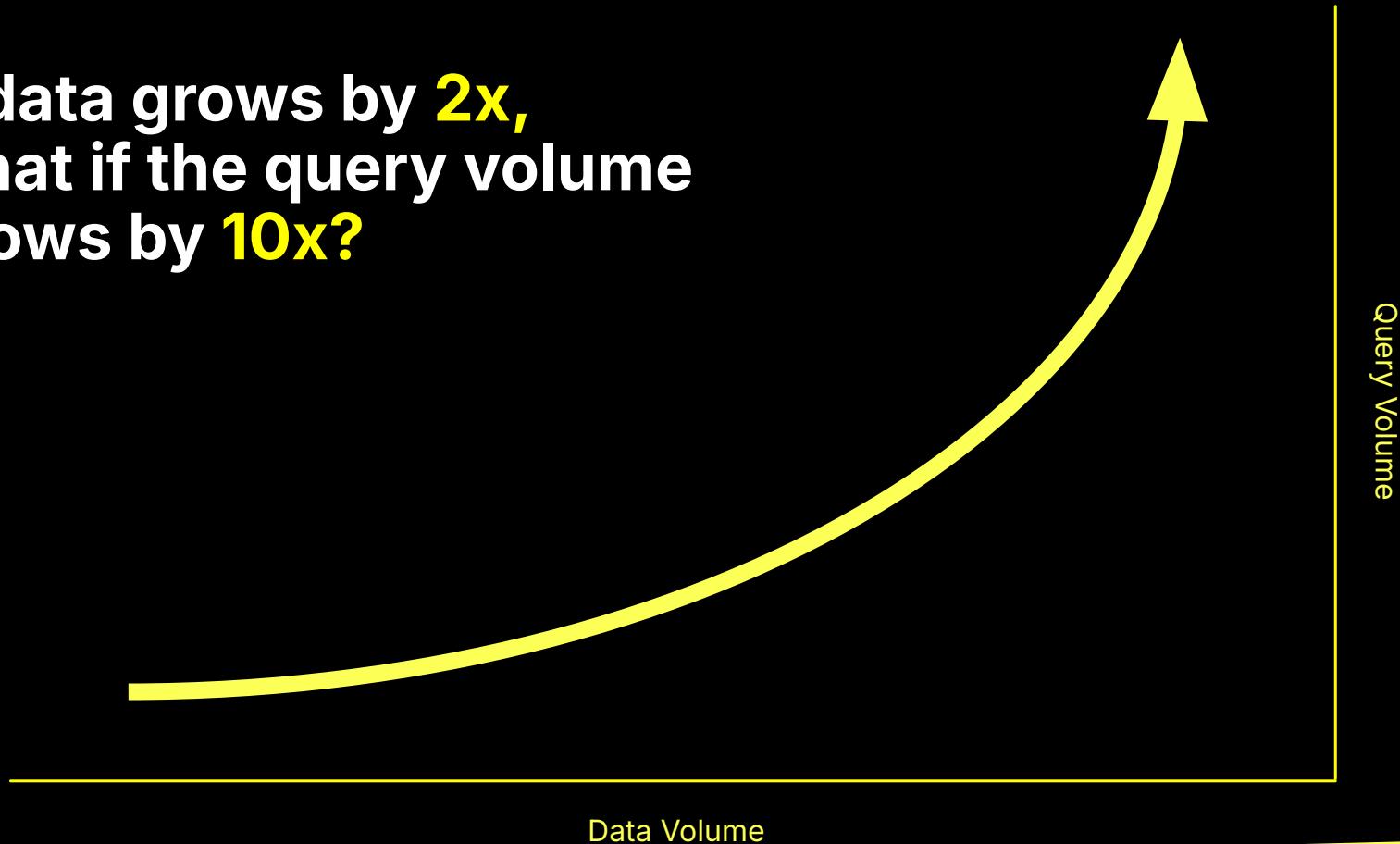


## ML and Gen AI

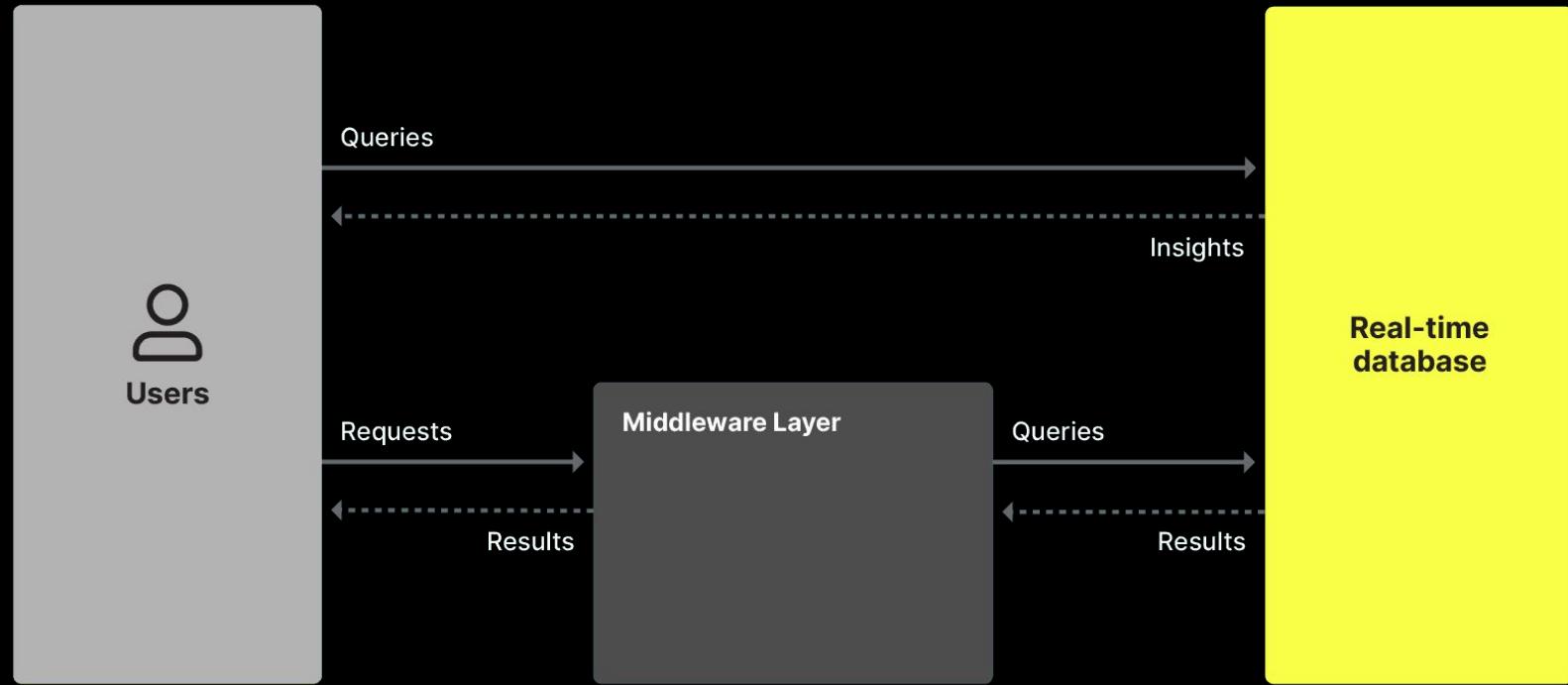
Execute fast and efficient vector search. Plug-and-play Generative AI models from any provider. Use lightning-fast aggregations to power model training at petabyte scale.



If data grows by **2x**,  
what if the query volume  
grows by **10x**?



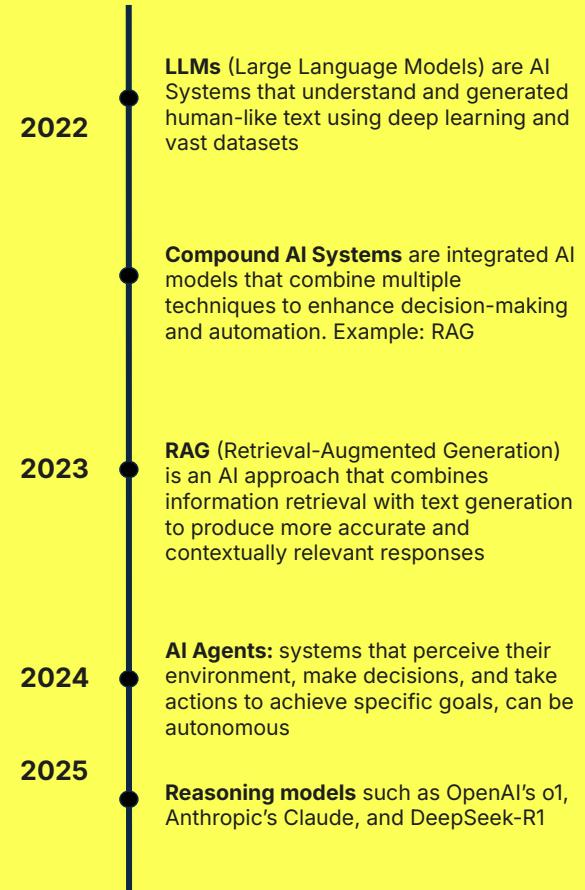
## Classic real-time analytics



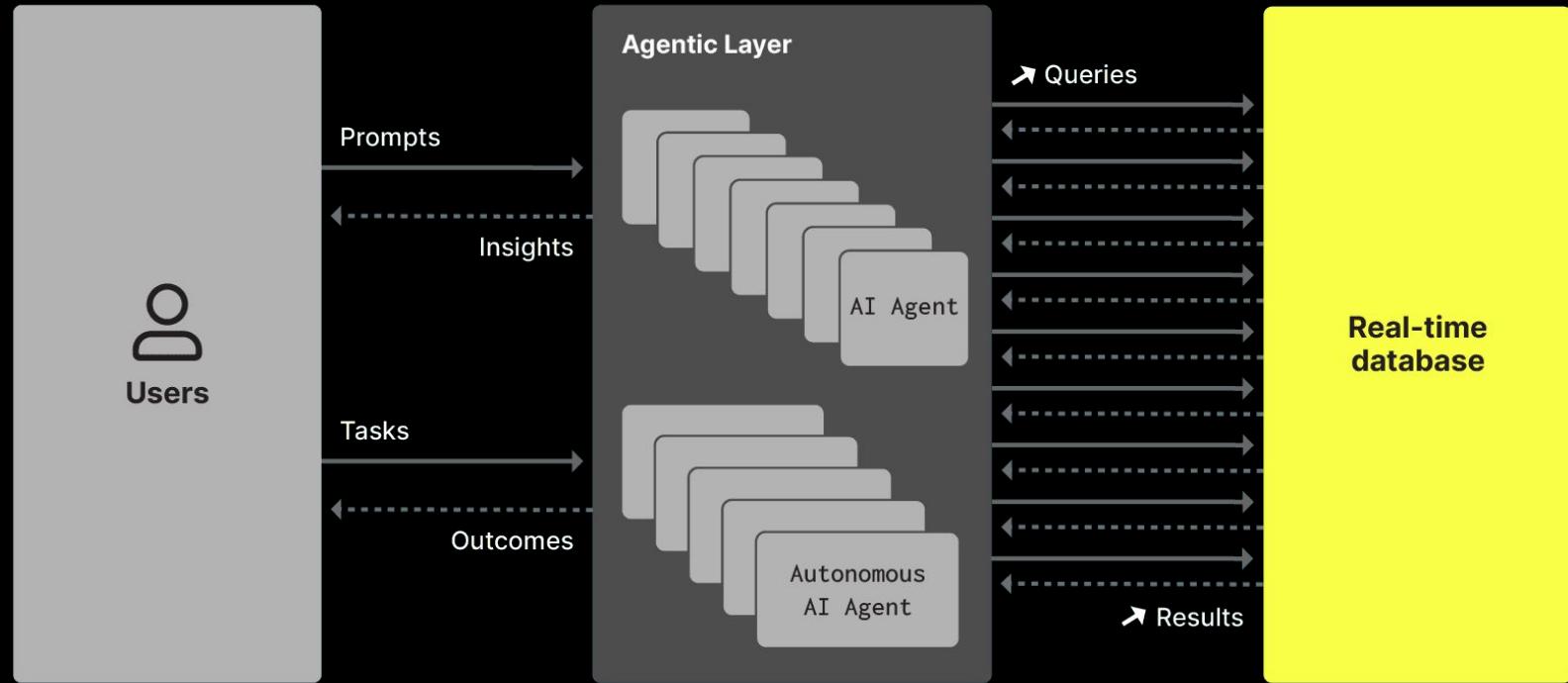
# The Emergence of Agents and Reasoning Models

## An agentic revolution

- Powered by LLMs, agents can understand, decide, and act
- Tools supercharge agents with real-world abilities
- Reasoning models introduce deliberate thinking before action
- AI agents are a new kind of “user” - tireless, autonomous, and prolific



## Agentic real-time analytics



# Why do you need real-time analytics for AI Agents?



## High-performance querying

Agents frequently issue multiple back-to-back queries when analyzing data or forming responses. Real-time databases provide the low-latency performance needed to support these explorative patterns without degrading user experience.



## Analytics in context

Real-time analytics databases are built for complex aggregations and fast pattern recognition across large datasets. This allows agents to go beyond raw data retrieval and surface meaningful business insights in real-time.



## Low-latency data

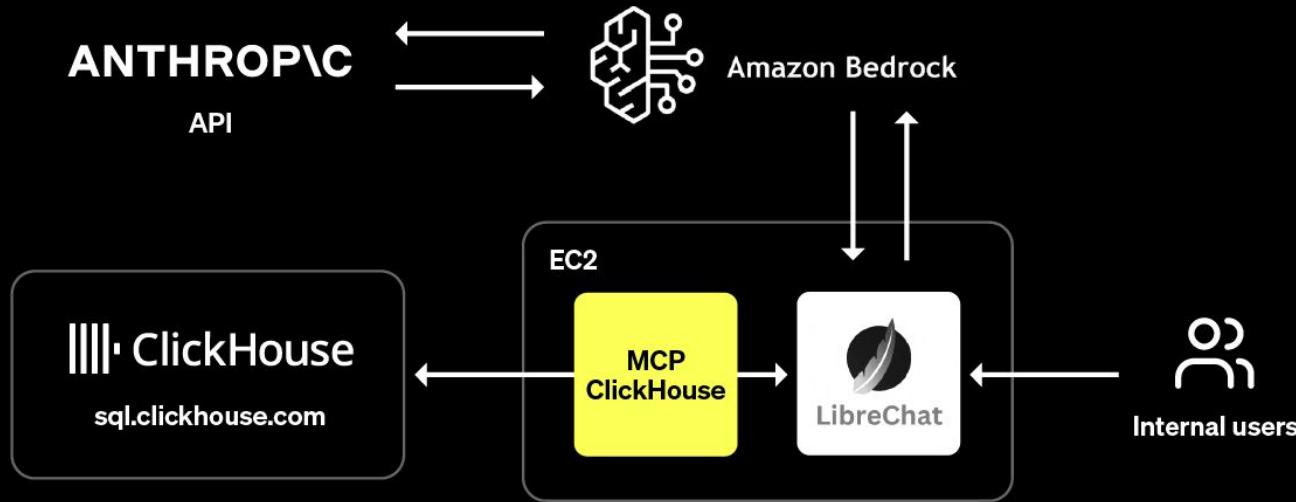
AI agents often operate autonomously and require the most recent data to respond accurately. Real-time databases ensure agents have immediate access to up-to-date information, allowing them to act on current trends and changing business conditions.



## Unified data access

AI agents need broad context to be effective. Real-time databases consolidate domain-specific data from multiple systems into one location, giving agents a reliable, consistent source of truth that's decoupled from transactional systems but deeply analytical in nature.

# Example architecture



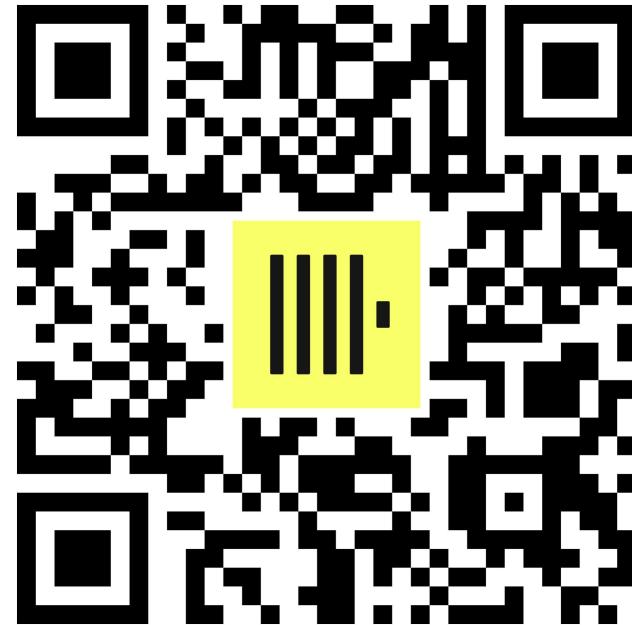
# Demo

“

Itching to try  
**AgentHouse**  
yourself?



Scan the code  
or visit [lm.clickhouse.com](https://lm.clickhouse.com)





## Connect with ClickHouse



slack  
from Salesforce



X



GitHub



ClickHouse  
Academy



### Try ClickHouse for your use case

- ClickHouse Cloud
- Download open source

### Learn

- Academy / certifications
- Blogs / YouTube

### Engage with our community

- Community Slack
- Monthly Community calls
- Meetups / events

**We are Hiring. Come Work with Us!**

# Get started today!

```
$ curl https://clickhouse.com/ | sh
```