Agentic AI

Anupam Datta (Snowflake)
John Mitchell (Stanford)

Reminder: Project pitches

- Sign up by Monday
- Present in class next week
- More details in class Thursday

Who are the founders of TruEra?

2023

TruEra was founded in 2014 by Anupam Datta, Somesh Jha, and Tomer Geva. Some key details about the founders:

- Anupam Datta - Professor of Computer Science and Electrical and Computer Engineering at Carnegie Mellon University. He served as TruEra's President and Chief Scientist.
- Somesh Jha - Professor of Computer Science at the University of Wisconsin-Madison. He was the Vice President of Research at TruEra.
- Tomer Geva - Previously worked at Microsoft Research and VMware. He was TruEra's CEO.

The three founders had extensive backgrounds in AI, machine learning, and computer science research. Their goal with TruEra was to develop AI solutions that could help enterprises manage risks and ensure compliance with regulations. The company was headquartered in Redwood City, California.

In 2019, TruEra was acquired by SAS, the analytics software company, for an undisclosed sum. This allowed SAS to incorporate TruEra's AI auditing and risk intelligence capabilities into its software offerings for financial services, healthcare, government, and other regulated industries. While the TruEra brand was absorbed into SAS, the founders

2025

A Who are the founders of TruEra?

TruEra founders 10 results

The founders of TruEra are:

Anupam Datta - Co-founder, President, and Chief Scientist. He was previously a professor at Carnegie Mellon University for 15 years, most recently as Professor and Director of the Accountable Systems Lab. Snowflake to Acquire Truer...

Shayak Sen - Co-founder and Chief Technology Officer. Snowflake to Acquire Truer...

Will Uppington - Co-founder and Chief Executive Officer. He previously held the role of Head of Product at Bloomreach. Snowflake to Acquire Truer...

TruEra was acquired by Snowflake in 2024, and all three co-founders joined Snowflake as part of the acquisition to continue developing AI observability capabilities for Snowflake's AI Data Cloud. Snowflake TruEra The company was founded around 2019 and specialized in AI quality management solutions that test, analyze, and monitor machine learning models.

Retry

Claude can make mistakes. Please double-check cited sources.

Reply to Claude...

Research Claude Sonnet 4

From Prompting LLMs to Agentic AI

Insights, Design Patterns & Evaluation

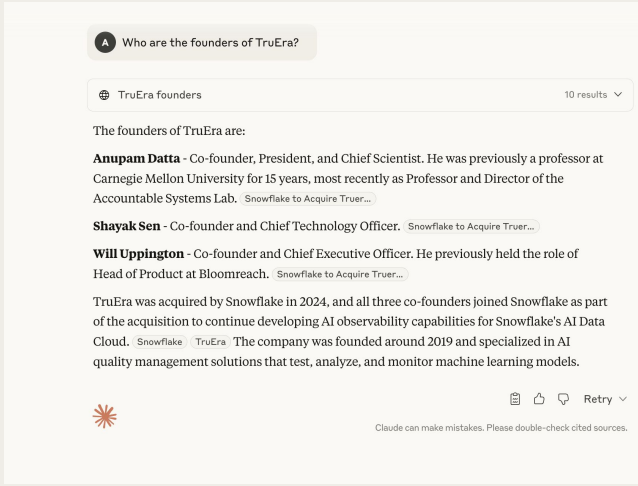
TruEra was founded in 2014 by Anupam Datta, Somesh Jha, and Tomer Geva. Some key details about the founders:

- Anupam Datta - Professor of Computer Science and Electrical and Computer Engineering at Carnegie Mellon University. He served as TruEra's President and Chief Scientist.
- Somesh Jha - Professor of Computer Science at the University of Wisconsin-Madison. He was the Vice President of Research at TruEra.
- Tomer Geva - Previously worked at Microsoft Research and VMware. He was TruEra's CEO.

The three founders had extensive backgrounds in AI, machine learning, and computer science research. Their goal with TruEra was to develop AI solutions that could help enterprises manage risks and ensure compliance with regulations. The company was headquartered in Redwood City, California.

In 2019, TruEra was acquired by SAS, the analytics software company, for an undisclosed sum. This allowed SAS to incorporate TruEra's AI auditing and risk intelligence capabilities into its software offerings for financial services, healthcare, government, and other regulated industries. While the TruEra brand was absorbed into SAS, the founders

2023



2025

Agentic Systems

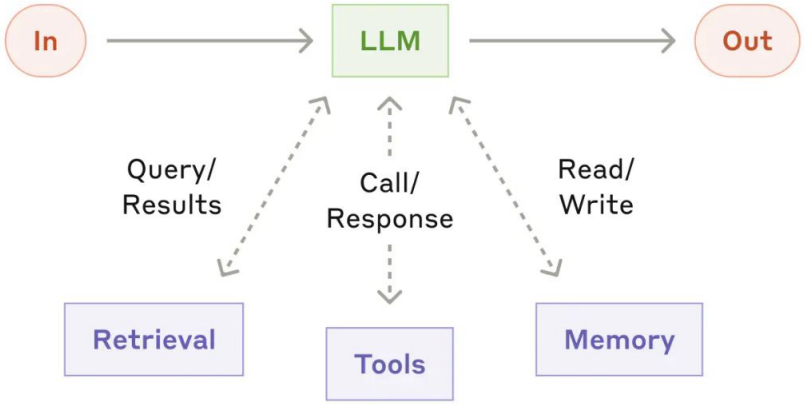
What are agents?

"Agent" can be defined in several ways. Some customers define agents as fully autonomous systems that operate independently over extended periods, using various tools to accomplish complex tasks. Others use the term to describe more prescriptive implementations that follow predefined workflows. At Anthropic, we categorize all these variations as agentic systems, but draw an important architectural distinction between workflows and agents:

- Workflows are systems where LLMs and tools are orchestrated through predefined code paths.
- Agents, on the other hand, are systems where LLMs dynamically direct their own processes and tool usage, maintaining control over how they accomplish tasks.

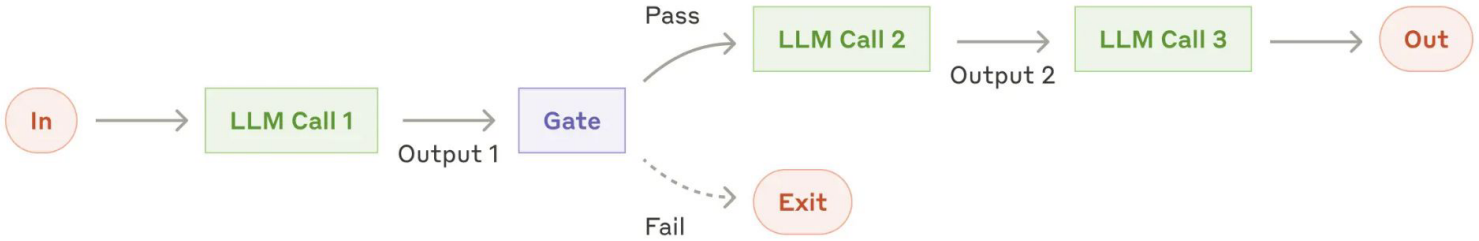
Source: Building effective agents

Building Block: The Augmented LLM



Source: Building effective agents

Workflow: Prompt Chaining

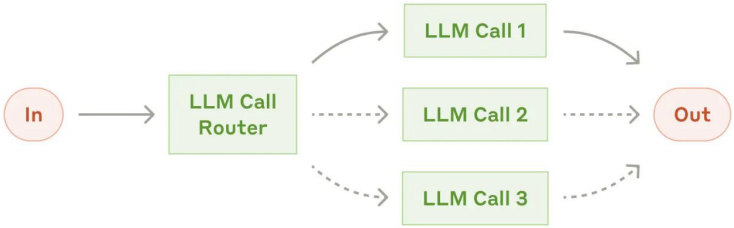


Examples

- Generating Marketing copy, then translating it into a different language
- Writing an outline of a document, checking that the outline meets certain criteria, then writing the document based on the outline.

Source: [Building effective agents]

Workflow: Routing

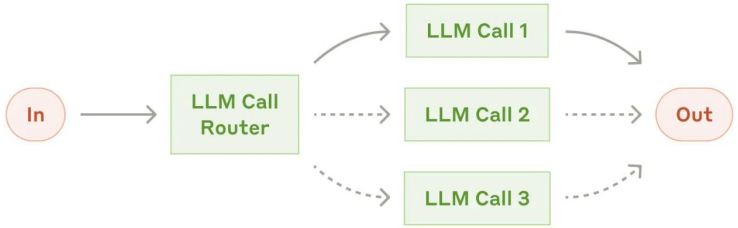


Examples

- Directing different types of customer service queries (general questions, refund requests, technical support) into different downstream processes, prompts, and tools.
- Routing easy/common questions to smaller models like Claude 3.5 Haiku and hard/unusual questions to more capable models like Claude 3.5 Sonnet to optimize cost and speed.

Source: Building effective agents

Workflow: Parallelization (1/2)

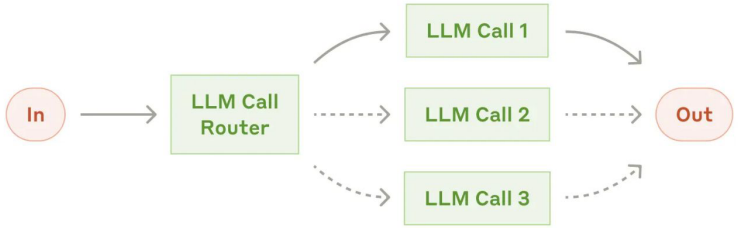


Examples: Sectioning

- Implementing guardrails where one model instance processes user queries while another screens them for inappropriate content or requests. This tends to perform better than having the same LLM call handle both guardrails and the core response.
- Automating evals for evaluating LLM performance, where each LLM call evaluates a different aspect of the model's performance on a given prompt.

Source: [Building effective agents]

Workflow: Parallelization (2/2)



Examples: Voting

- Reviewing a piece of code for vulnerabilities, where several different prompts review and flag the code if they find a problem.
- Evaluating whether a given piece of content is inappropriate, with multiple prompts evaluating different aspects or requiring different vote thresholds to balance false positives and negatives.

Source: Building effective agents

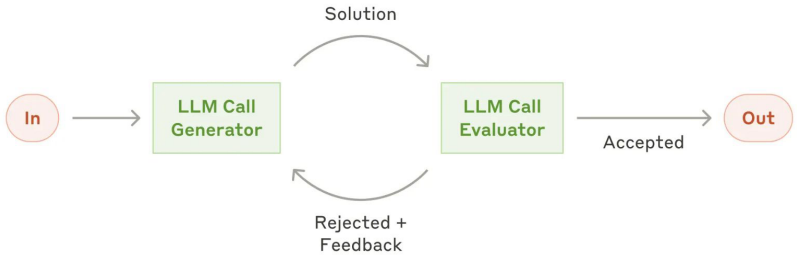
Workflow: Orchestrator-Workers

Examples

- Coding products that make complex changes to multiple files each time.
- Search tasks that involve gathering and analyzing information from multiple sources for possible relevant information.

Source: [Building effective agents]

Workflow: Evaluator-Optimizer

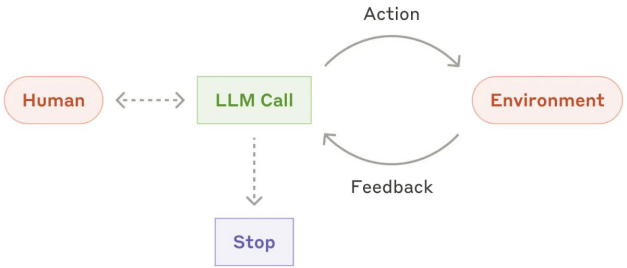


Examples

- Literary translation where there are nuances that the translator LLM might not capture initially, but where an evaluator LLM can provide useful critiques.
- Complex search tasks that require multiple rounds of searching and analysis to gather comprehensive information, where the evaluator decides whether further searches are warranted.

Source: [Building effective agents]

Workflow: Agents



Examples

- A coding Agent to resolve SWE-bench tasks, which involve edits to many files based on a task description;
- Anthropic “computer use” reference implementation, where Claude uses a computer to accomplish tasks.

Source: Building effective agents

Agent Building Blocks

- Planning
- Tool Use (incl. web search)
- Reflection
- Memory
- Multi agent collaboration

Agent Building Blocks

References

Reflection

- Self-Refine: Iterative Refinement with Self-Feedback
- Reflexion: Language Agents with Verbal Reinforcement Learning

Tool use

- Gorilla: Large Language Model Connected with Massive APIs
  - Gorilla: Large Language Models Connected with Massive APIs
  - Gorilla LLM: Teach LLMs to Use Tools at Scale
  - Evals: Berkeley Function-Calling Leaderboard

Planning

- Chain-of-Thought Prompting Elicits Reasoning in Large Language Models
- HuggingGPT: Solving AI Tasks with ChatGPT and its Friends in Hugging Face

Multi agent collaboration

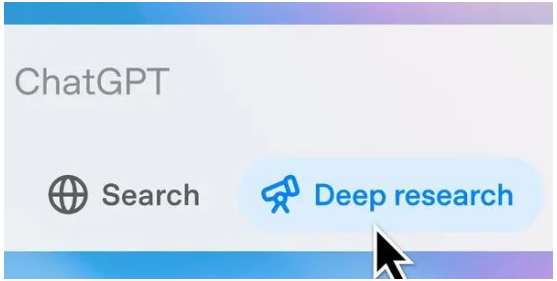
- ChatDev: Communicative Agents for Software Development
- AutoGen: Enabling Next-Gen LLM Applications via Multi-Agent Conversation

Data Agents

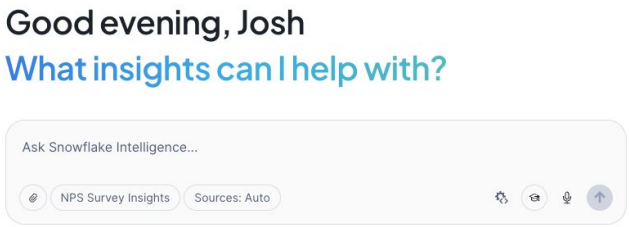
A data agent is an autonomous or semi-autonomous system that:

- Connects to data sources (databases, APIs, files, streams, sensors, Web etc.)
- Understands queries expressed in natural language or code
- Plans & performs actions such as query decomposition, data retrieval, analysis or visualization
- Provides insights or takes decisions based on the data

Data Agents Gaining Widespread Adoption

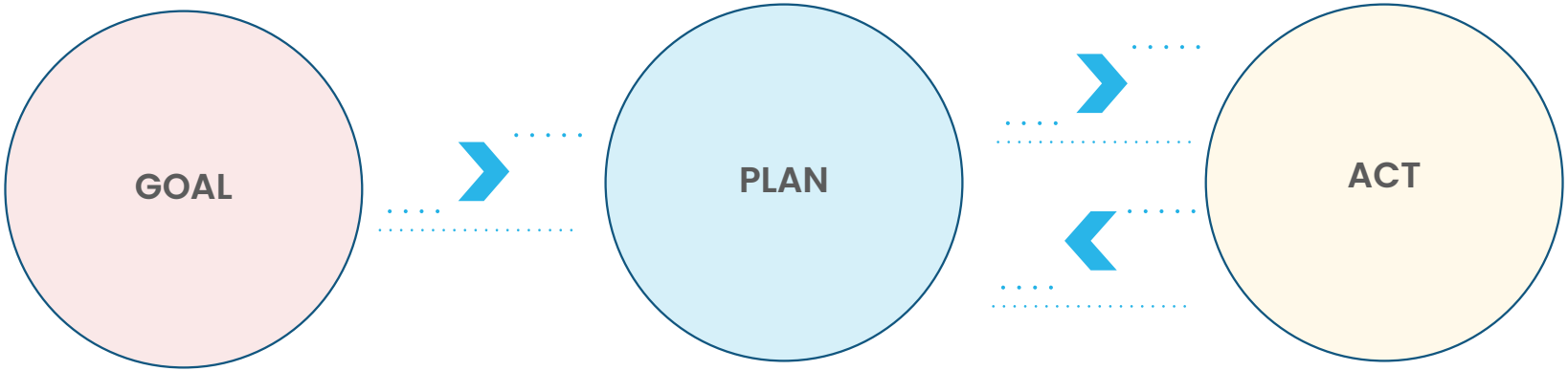


Deep Research



Snowflake Intelligence

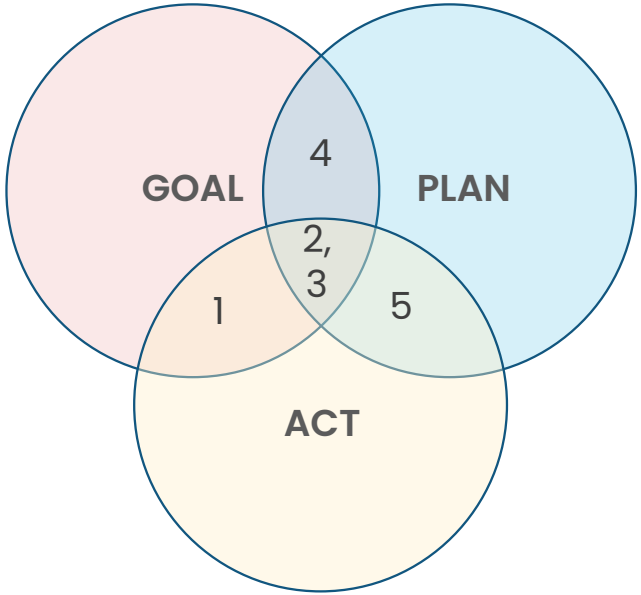
How Do Agents Work?



How Do Agents Work *Well*?

Trustworthy Agents Execute with **G**oals, **P**lans and **A**ctions Aligned.

What is Your Agent's GPA?



1. Goal Fulfillment

1A. Answer Relevance

2. Logical Consistency

3. Execution Efficiency

4. Plan Quality

4A. Tool Selection

5. Plan Adherence

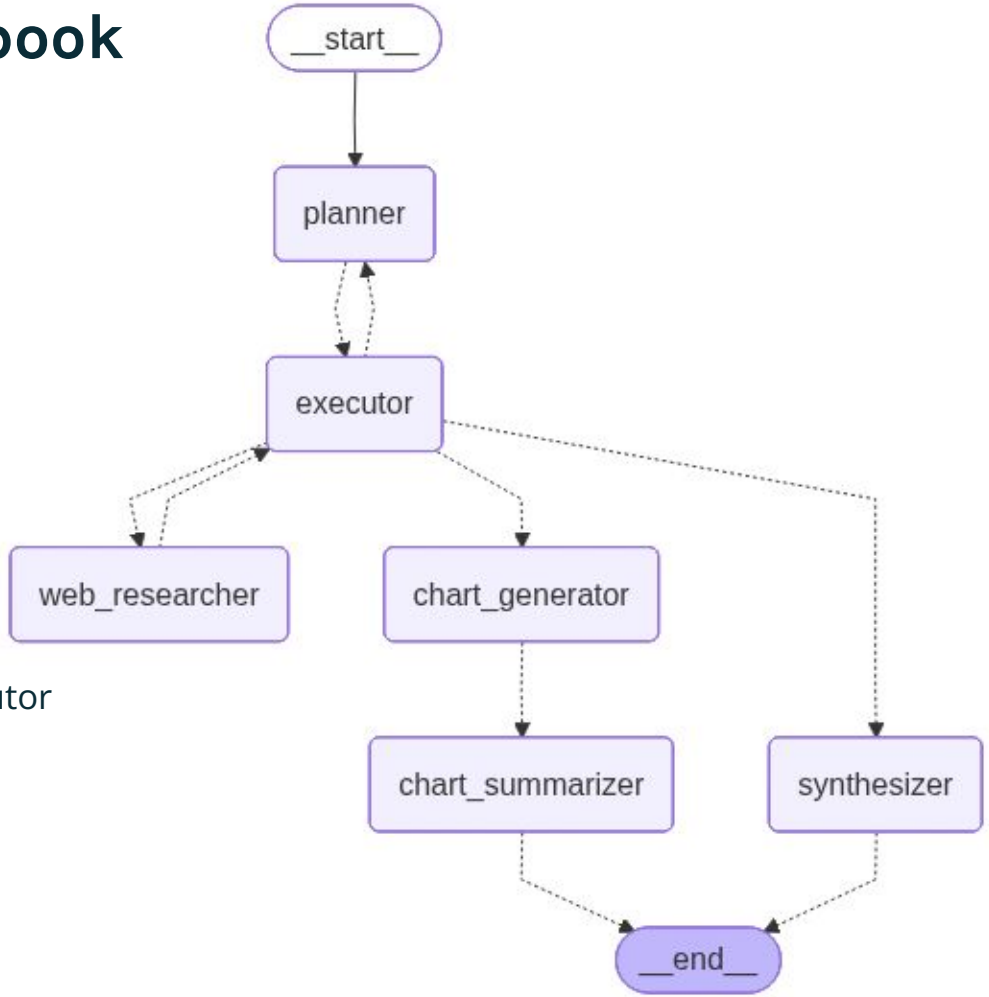
5A. Tool Calling

Homework 1: Build & Evaluate a Data Agent

- Build with LangGraph
  - Background: LangGraph course
- Evaluate with TruLens
- Walkthrough of Homework 1 notebook

LangGraph Agents in Notebook

- Memory
  - Augmented LangGraph State
- Reflection
  - Performed by Executor/Planner
- Tool Use
  - Web Researcher uses Tavily Tool
  - Charting uses Python Tool
- Planning
  - Performed by Planner
- Multi agent collaboration
  - Coordinated through Planner and Executor



Agent Memory for the Example Notebook

- Global memory
  - User query
  - Available agents
  - Current Plan
  - Replan tracking

- Short Term Memory
  - Current step
  - Agent query
  - Agent selection reason

Who are the founders of TruEra?

2023

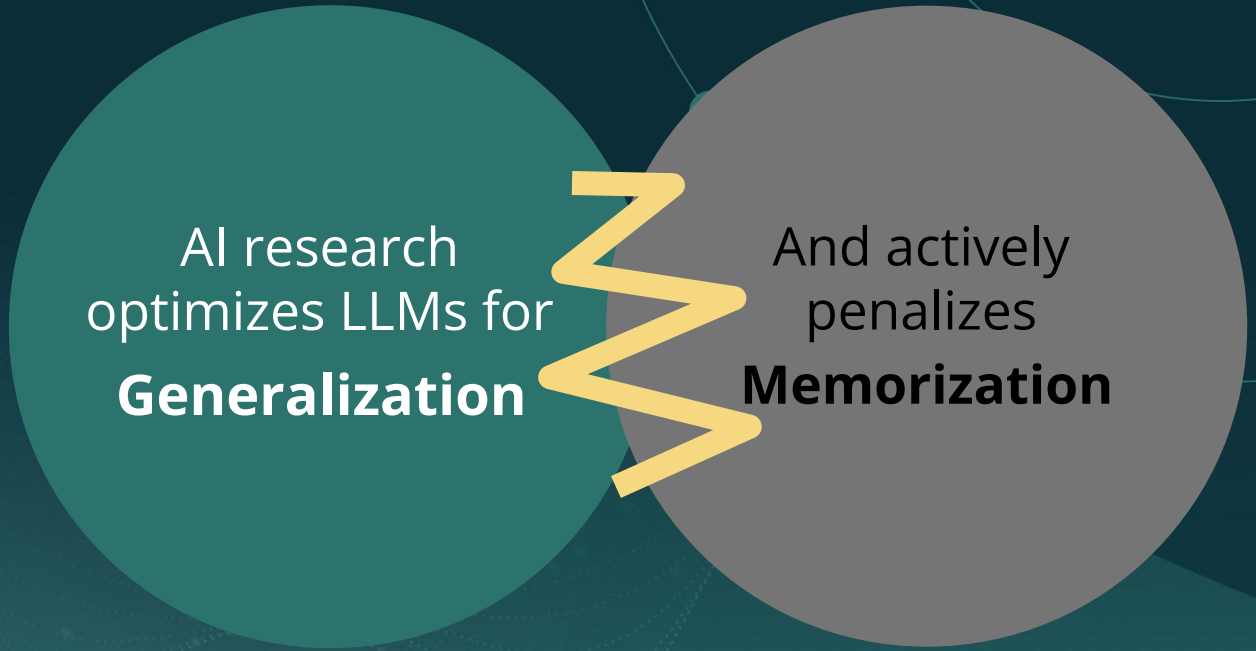
TruEra was founded in 2014 by Anupam Datta, Somesh Jha, and Tomer Geva. Some key details about the founders:

- Anupam Datta - Professor of Computer Science and Electrical and Computer Engineering at Carnegie Mellon University. He served as TruEra's President and Chief Scientist.
- Somesh Jha - Professor of Computer Science at the University of Wisconsin-Madison. He was the Vice President of Research at TruEra.
- Tomer Geva - Previously worked at Microsoft Research and VMware. He was TruEra's CEO.

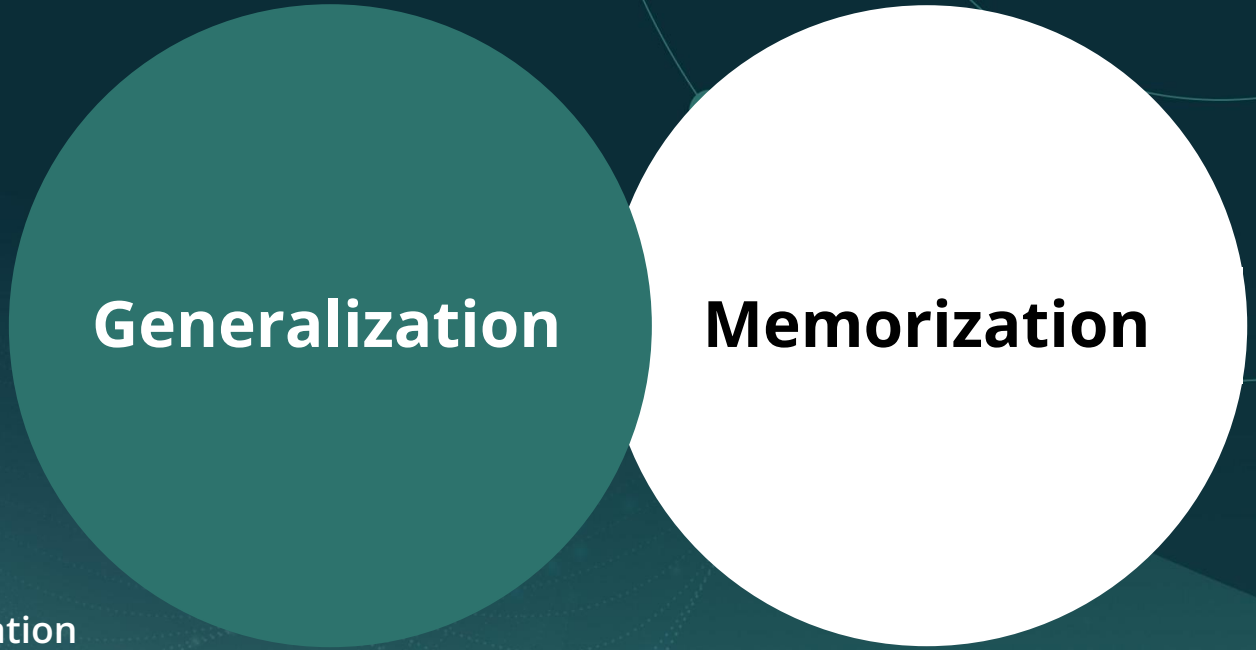
The three founders had extensive backgrounds in AI, machine learning, and computer science research. Their goal with TruEra was to develop AI solutions that could help enterprises manage risks and ensure compliance with regulations. The company was headquartered in Redwood City, California.

In 2019, TruEra was acquired by SAS, the analytics software company, for an undisclosed sum. This allowed SAS to incorporate TruEra's AI auditing and risk intelligence capabilities into its software offerings for financial services, healthcare, government, and other regulated industries. While the TruEra brand was absorbed into SAS, the founders

This overlap is really murky



Focus LLMs on ‘General’ Tasks

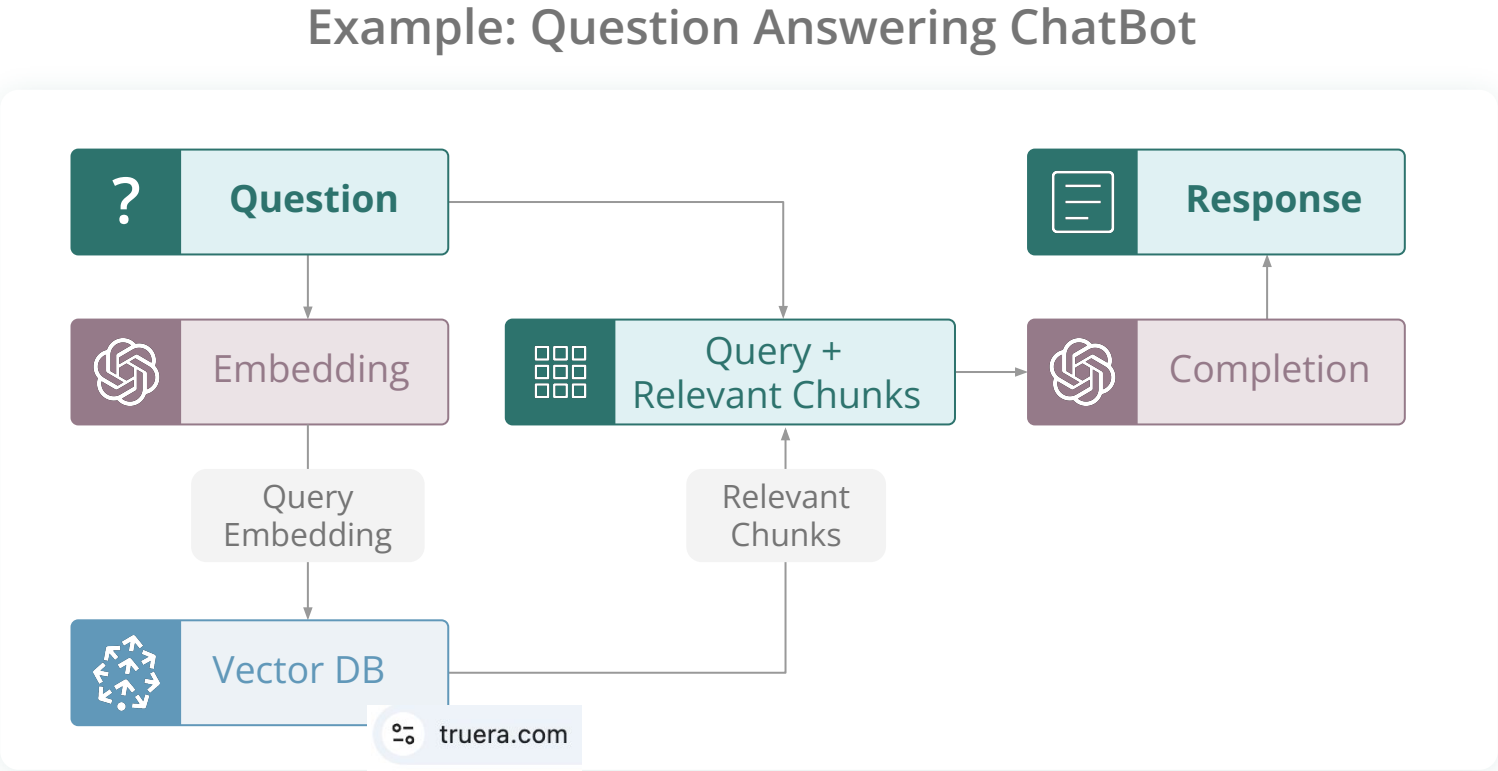


- ✓ Summarization
- ✓ Text Embedding
- ✓ Logical Inference
- ✓ Planning

Leave memorization
to something else

LLMs Need a Knowledge Source

Enter Retrieval Augmented Generation (RAGs)



How RAGs work

1. Indexing
2. Retrieval
3. Augmentation
4. Generation

How RAGs work: Indexing

The custom data source, which can include documents, articles, databases, and more, is prepared for retrieval.

- The documents are first broken down into smaller, manageable "chunks" to fit within the LLM's context window during generation.
- An embedding model then converts each chunk of text into a numerical representation called a vector embedding. Vector embeddings capture the semantic meaning of the text, so similar content will have vectors that are numerically "close" to one another in a multi-dimensional space.
- All of these vector embeddings are stored in a specialized database, called a vector database.
- **Metadata Association:** Each vector is typically stored with an ID that links it back to its original chunk and stores additional metadata, such as the source document, date, or other relevant information.

How RAGs work: Retrieval

When a user submits a query, the system retrieves the most relevant information from the vector database.

- The user's query is converted into a vector embedding using the same embedding model that processed the source documents.
- The system then performs a vector search to find the most similar vectors in the database based on their proximity to the query's vector.
- The document chunks corresponding to the top-ranking vectors are retrieved

How RAGs work: Augmentation

The retrieved information is used to "augment" the original user query.

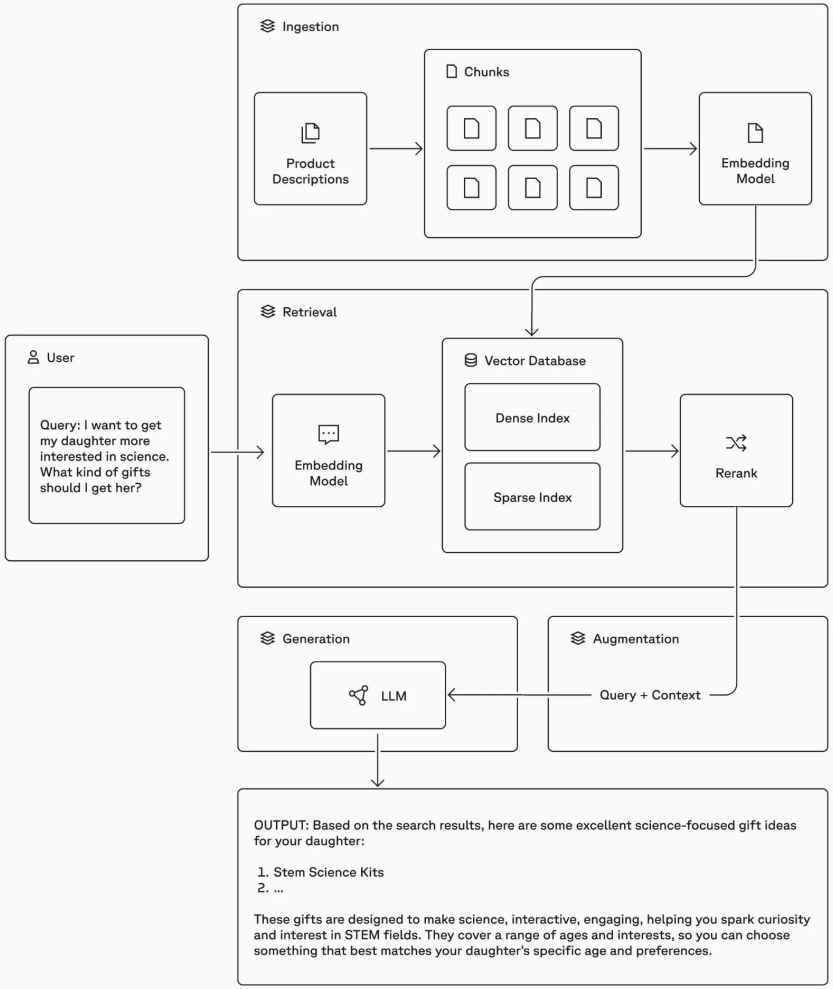
- The system creates a new, enhanced prompt for the LLM that includes the user's original question along with the retrieved, contextually relevant document chunks.
- This process, sometimes called "prompt stuffing," provides the LLM with the specific, factual information it needs to construct an accurate response.

How RAGs work: Generation

The augmented prompt is sent to the LLM, which uses the new information to create a final, detailed, and context-aware response.

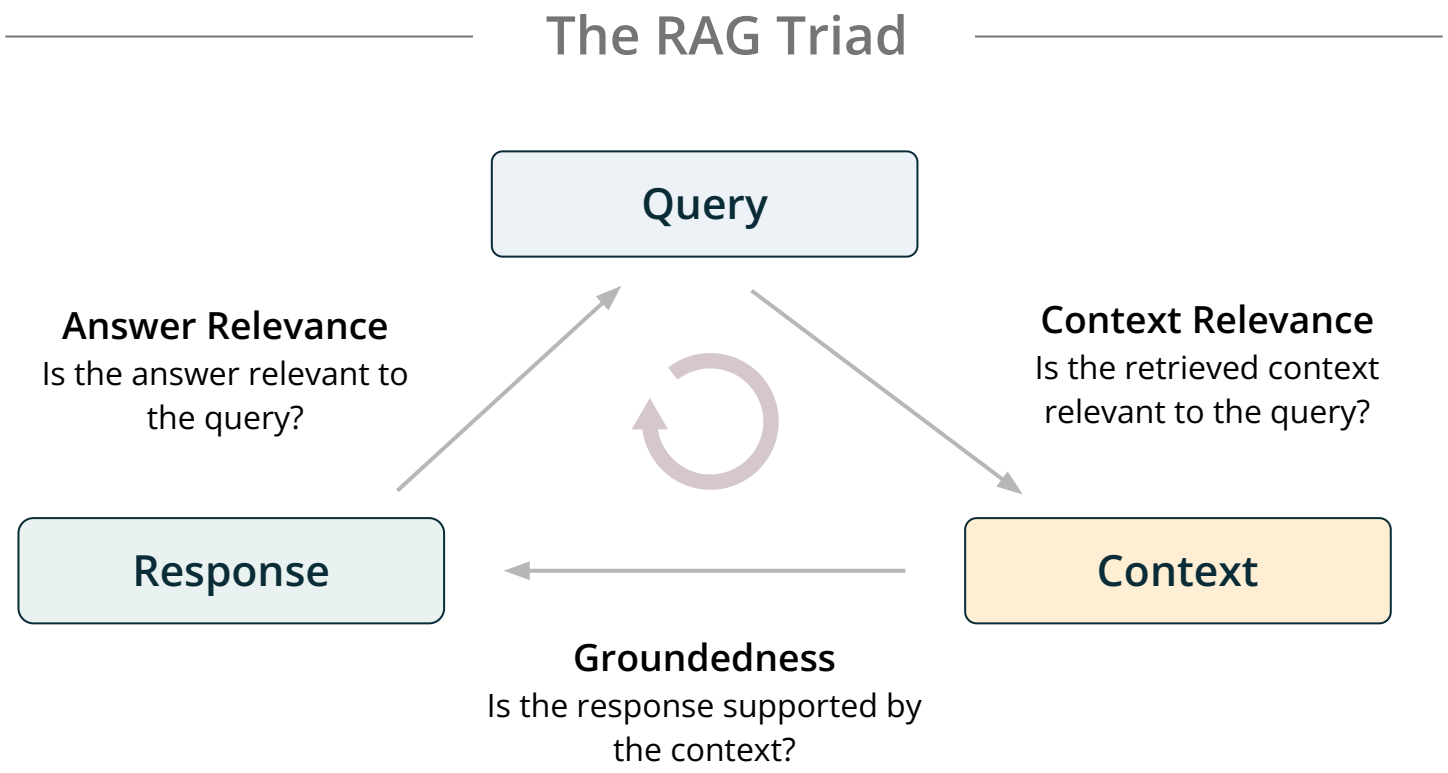
- The LLM synthesizes its own general knowledge with the retrieved data.
- A key benefit is that RAG systems can also provide source citations, allowing the user to verify the information and build trust.

Hybrid RAG

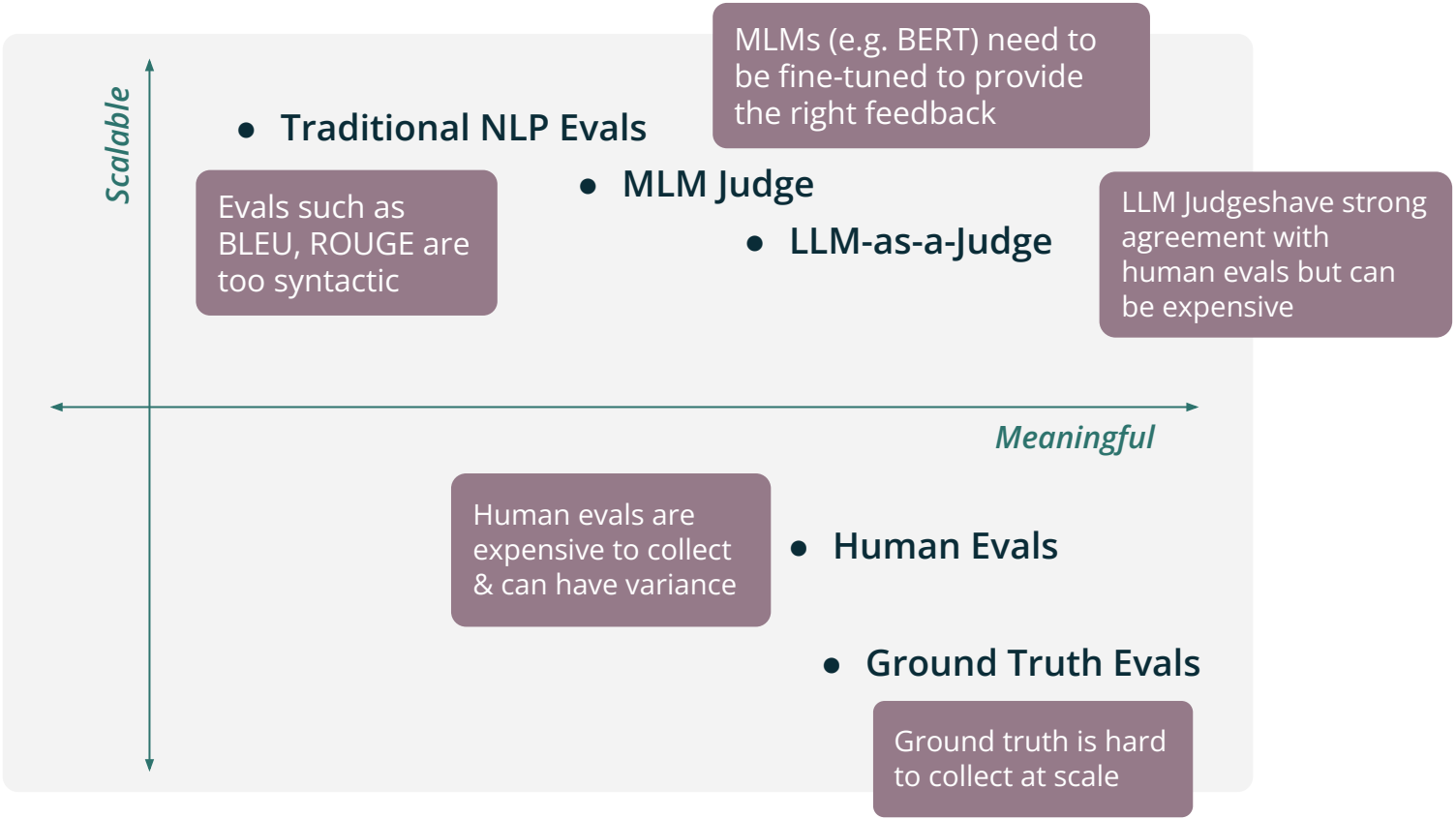


Cortex Search, Pinecone

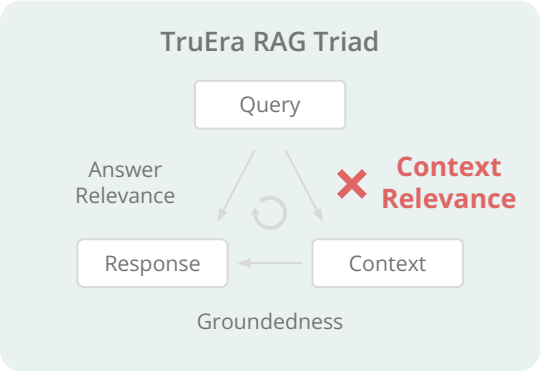
Evaluating RAGs



Choosing the Right Evals



Retrieval Failure



Input

Who is Shayak?

Response

Shayak is a computer scientist who obtained his PhD in Computer Science from Carnegie Mellon University and BTech in Computer Science from the Indian Institute of Technology, Delhi. He has been building systems and leading research to make machine learning and big data systems more explainable, privacy compliant, and fair for the past 10 years. He is also a member of the Bank of England's AI Public-Private Forum and the OECD Global Partnership on AI.

Feedback

qs_relevance = 0.55

	question	statement	result
0	Who is Shayak?	When Shayak started building production grade machine learning models for algorith	1.0
1	Who is Shayak?	When Shayak started building production grade machine learning models for algorith	1.0
2	Who is Shayak?	Most recently, Shameek was Group Chief Data Officer at Standard Chartered Bank, wh	0.1
3	Who is Shayak?	Shameek has spent most of his career in driving responsible adoption of data analyti	0.1

Lack of Groundedness

Selected LLM Application: Summarize_v1

Selected Record ID: record_hash_b63e7be3dac4d43a3d79178e22710b4c

Input [Select.RecordInput]

#Person1#: Hello, Is this room service? #Person2#: Yes, May I help you? #Person1#: This is the room 1425, we asked for the room service an hour ago. #Person2#: we are very sorry to cause you a lot of inconvenience. #Person1#: What's the matter? #Person2#: We're rather busy right now. It will take another 15minutes #Person1#: Is it really going to take that long, will you rush the order? #Person2#: I'm afraid it will take 15 minutes at most. #Person1#: ha, well, we have no choice.

Response [Select.RecordOutput]

Room 1425 called room service an hour ago and they were told it will take an additional 15 minutes for their order to arrive. They are not happy about the wait but have no other option.

groundedness_measure = 0.7

	statement	result	reason
0	rson1#: Th Room 1425 called room service an hour ago and they were told it will take an additio	0.	

Statement Sentence: Room 1425 called room service an hour ago and they were told it will take an additional 15 minutes for their order to arrive.,
Supporting Evidence: This is the room 1425, we asked for the room service an hour ago. We're rather busy right now. It will take another 15 minutes
Score: 10

Statement Sentence: They are not happy about the wait,
Supporting Evidence: NOTHING FOUND
Score: 1

Statement Sentence: but have no other option.,
Supporting Evidence: ha, well, we have no choice.
Score: 10

bert_score = None

No feedback details.

Answering the Wrong Question

Input

Which year was Hawaii's state song written?

Response

Hawai`i Pono`i

Feedback

qs_relevance = 0.325

	question	statement	result
0	Which year was Hawaii's state song written?	"Hawai`i Pono`i" is the state song of Hawaii. The words were written by King David K	1.0
1	Which year was Hawaii's state song written?	The American business people made Hawaii into a republic for a short time. The new	0.1
2	Which year was Hawaii's state song written?	Hawaii (sometimes spelled "Hawai'i".) is a U.S. state and the only U.S. State that is in	0.1
3	Which year was Hawaii's state song written?	1874 – Hawaii signs a treaty with the United States granting exclusive trade rights. 18	0.1

relevance = 0.1

	prompt	response	result
0	Which year was Hawaii's state song written?	Hawai`i Pono`i	0.1

Reminder: Project pitches

- Sign up by Sunday/Monday
- Present in class next week
- More details in class Thursday