Gartner Webinars

Gartner delivers actionable, objective insight, guidance and tools to enable stronger performance on your organization's mission critical priorities

Gartner®

Enhance your webinar experience

Ask a Question

Download Attachments

Watch Again

New Techniques for AI and GenAI

Connect with Gartner

Van Baker

VP Analyst

Gartner

The future of artificial intelligence is not about man versus machine, but rather man with machine. Together, we can achieve unimaginable heights of innovation and progress.

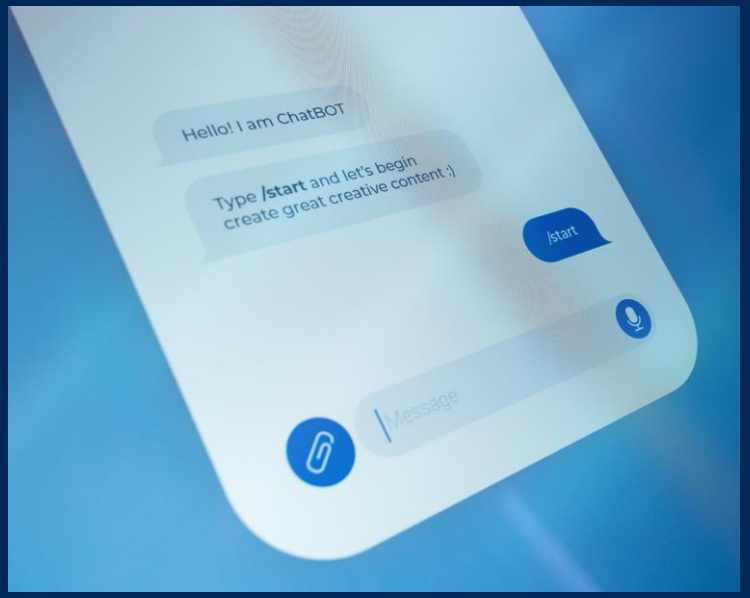
Fei-Fei Li

Co-director, Stanford Institute for Human-Centered AI

Gartner®

Three New Techniques Changing AI and GenAI

- GraphRAG vs RAG
- AI Agents
- Neuro-symbolic Models



GraphRAG vs. RAG

The Elements of RAG

Retriever

- Fetches relevant documents from an external knowledge base
- Typically utilizes vector representations of the text
- Can be a simple keyword-based search, semantic search, or vector search.

Augmentation

- The process of combining information from the retrieved documents with the input query.
- Done at the input layer by concatenating the query and retrieved documents
- Done at the intermediate layer by incorporating retrieved information into the model's hidden states
- Done at the output layer by post-processing the model's output.

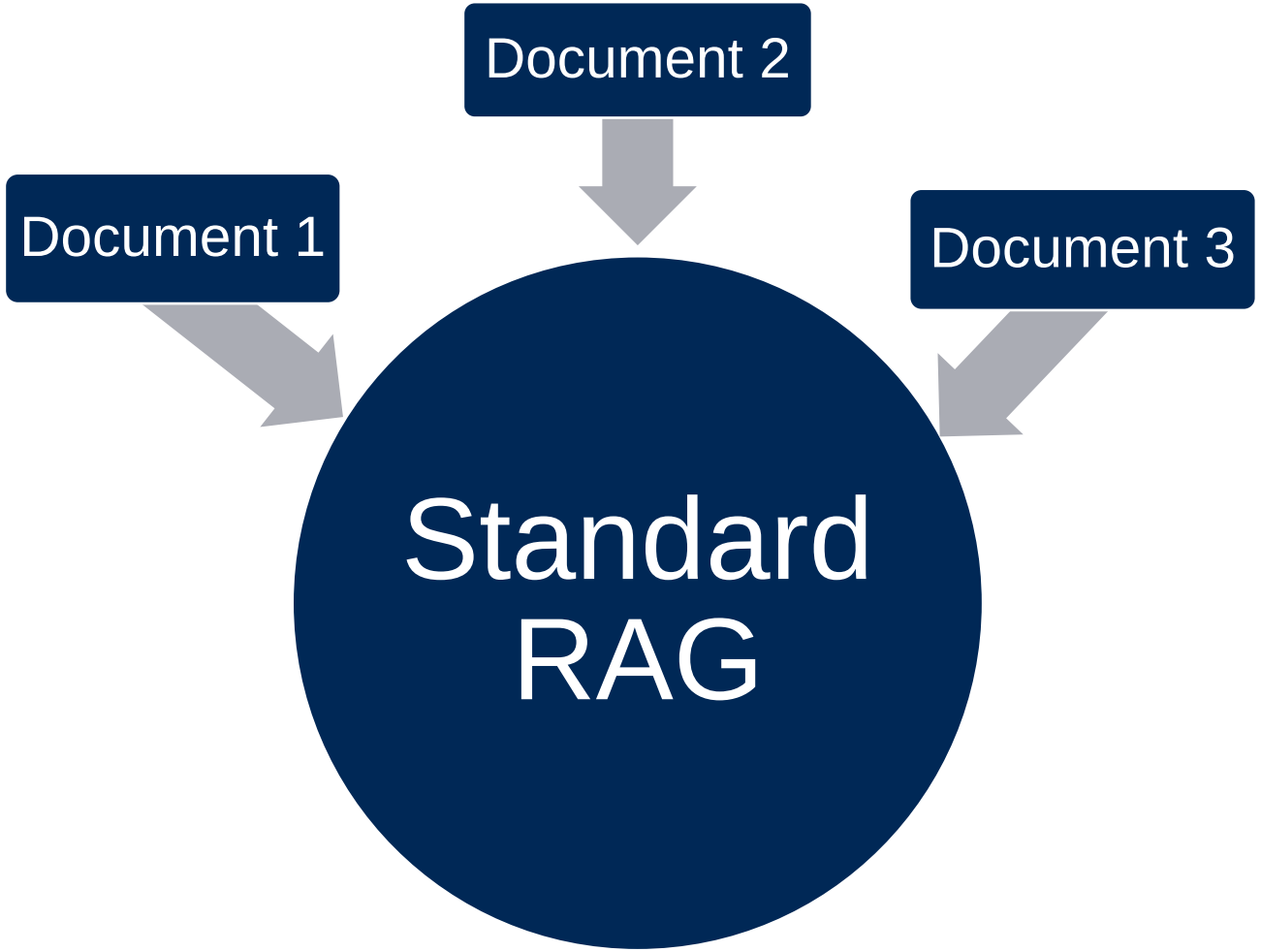
Generator

- LLM takes the retrieved documents and the original query as input.
- Generates text conditioned on the retrieved information to synthesize a response.
- The output of the generator is the final response of the RAG system.

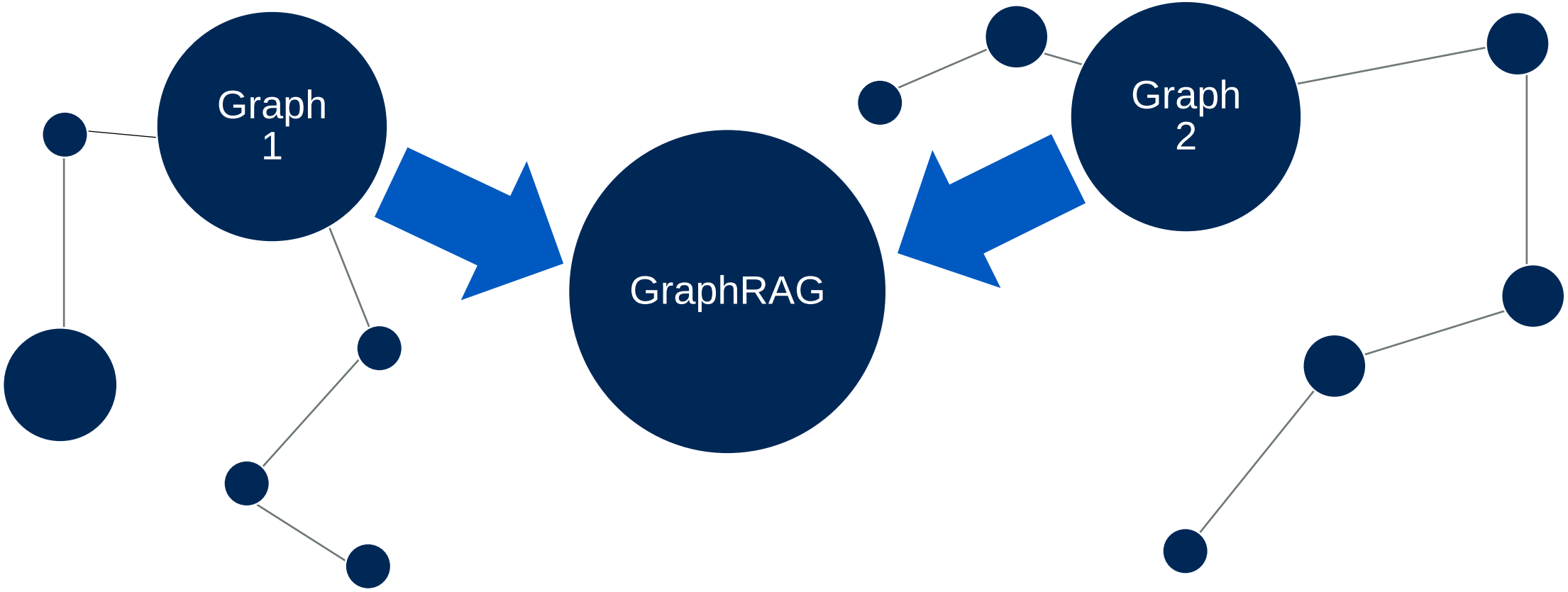
Three Types of Standard RAG

- Naïve
  - Follows a traditional "retrieve-read" framework.
  - Utilizes indexing, retrieval, and generation processes.
- Advanced
  - Extends Naive RAG with additional modules that enhance LLM performance
  - Query rewriting, chunk reranking, and prompt summarization.
- Modular
  - Allows addition, replacement, or rearrangement of modules within the RAG process.
  - Enables customization to address specific problem contexts.

Standard RAG



GraphRAG



Graph RAG vs. RAG

Feature	Standard RAG	GraphRAG
Retrieval Method	• Typically relies on vector similarity search in a vector database of text chunks.	• Leverages an LLM-derived knowledge graph to retrieve relevant information. • Index is built from the source documents and includes entities, relationships, and community summaries.
Target Queries	• Primarily focused on answering local questions about specific pieces of information within a document.	• Designed to handle both local (specific) and global (holistic) questions about a text corpus. • Excels at answering questions that require connecting disparate pieces of information or summarizing semantic concepts across a large collection of documents.
Strengths	• Effective for retrieving specific information based on keyword or semantic similarity. • Relatively simple to implement and integrate with LLMs.	• Improved comprehensiveness and diversity of answers, especially for global queries. • Ability to handle complex questions requiring multi-hop reasoning. - Provides source provenance for generated answers.
Limitations	• May struggle to answer questions requiring reasoning or synthesis of information.	• More complex to implement and requires a significant upfront investment to build the knowledge graph.

AI Agents

Market Confusion Around AI Agents

- Terminology is quite varied and results in confusion
- Technology providers are not helping *(Microsoft and Salesforce include assistants)*
- Various providers and clients believe:
  - Agents are models
  - Agents are assistants (Copilots)
  - Agents are not different than RPA
  - Agents are autonomous bots
- The last assumption is closest to the Gartner definition

Gartner Definition

AI agents are autonomous or semiautonomous software entities that use AI techniques to perceive, make decisions, take actions and achieve goals in their digital or physical environments.

Agent Expectations

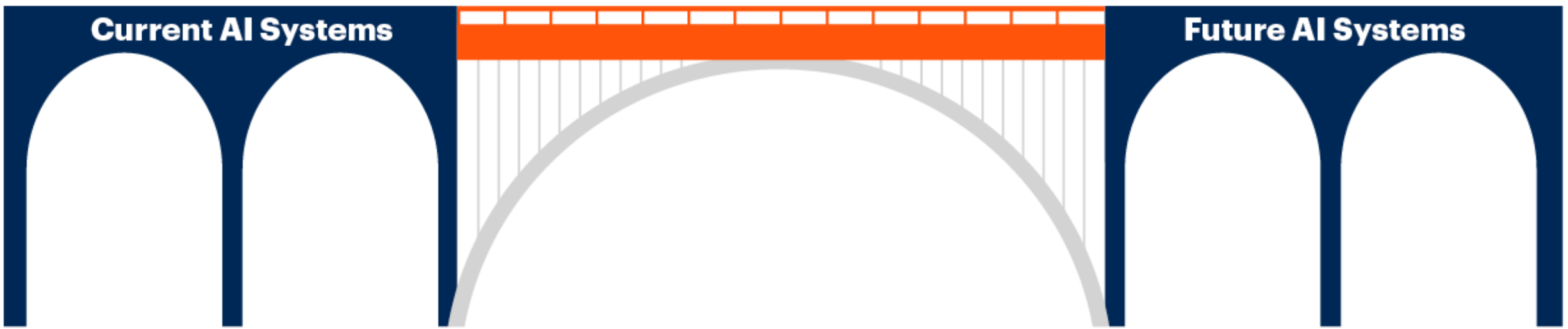
Closing the AI Agency Gap

Illustrative

- Require strong supervision
- Execute simple tasks
- Static behavior

AI Agents

- Act autonomously
- Execute complex goals
- Adaptable behaviour

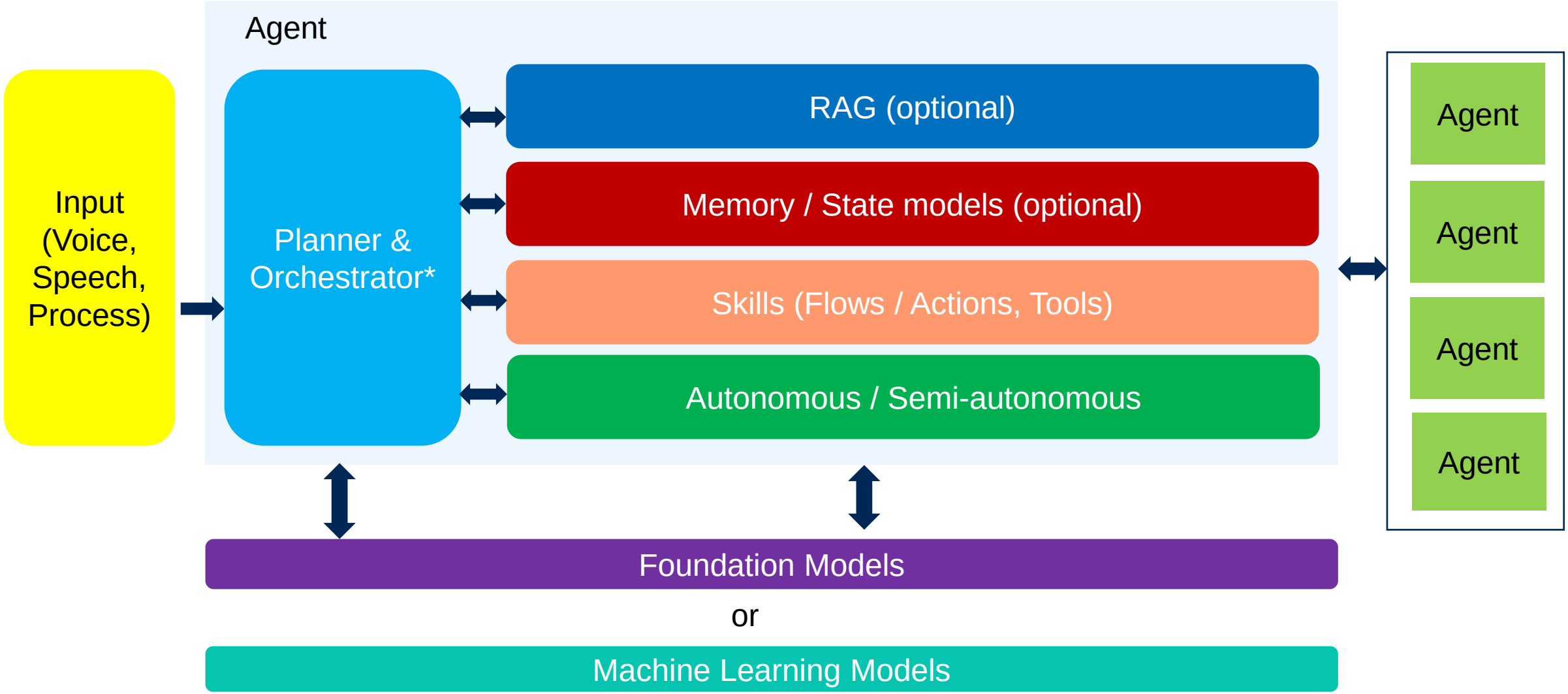


Source: Gartner

806843_C

Gartner

AI Agents Critical Elements



* Looping / Reflection

Potential Blockers for AI Agents

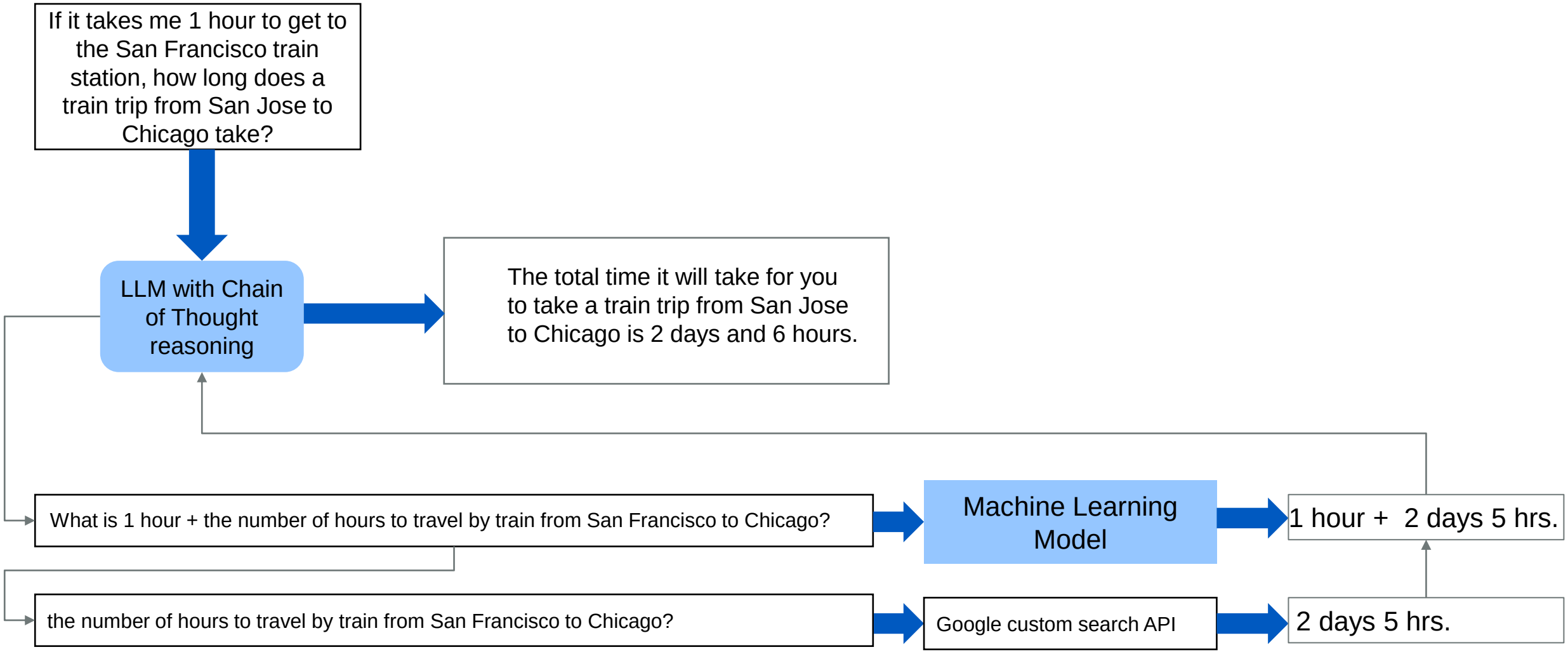
- Hallucinations
- Lack of monitoring / control tools
- Potential for serious damage
- Market confusion
- Workforce resistance
- Poor RAG implementation
- Poor prompt generation

Neurosymbolic AI

Neurosymbolic AI Definition

Neuro-symbolic AI is a form of composite AI that combines machine learning (ML) methods and symbolic systems (for example, knowledge graphs) to create more robust and trustworthy AI models. This fusion enables the combination of probabilistic models with explicitly defined rules and knowledge to give AI systems the ability to better represent, reason and generalize concepts. This approach provides a reasoning infrastructure for solving a wider range of business problems more effectively.

Neurosymbolic AI Example



Why Neurosymbolic AI

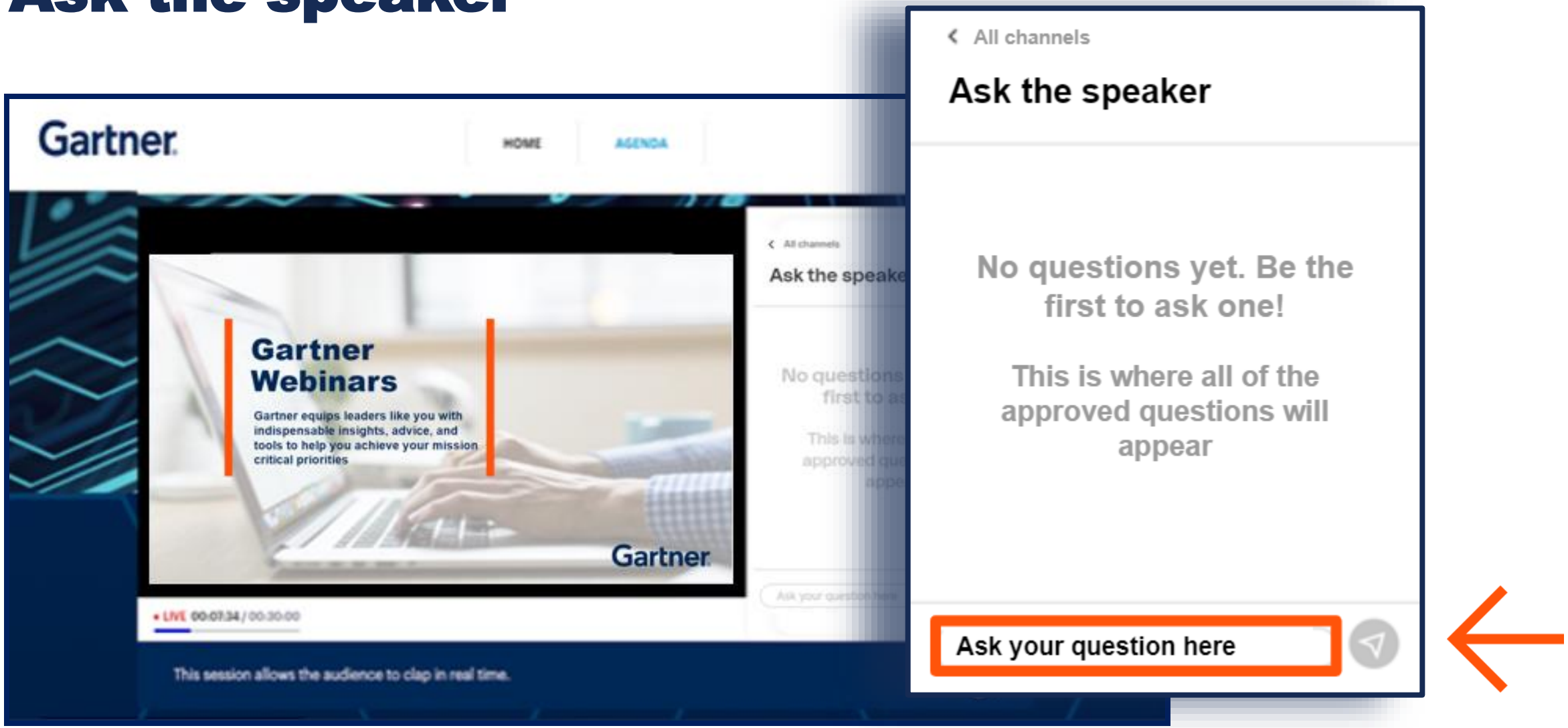
- Neuro-symbolic AI is important because it addresses limitations in current AI systems, such as incorrect outputs, lack of generalization to a variety of tasks and an inability to explain the steps that led to an output.
- Neuro-symbolic AI will have an impact on the efficiency, adaptability and reliability of AI systems used across business processes. The integration of logic and multiple reasoning mechanisms brings down the need for ever larger AI models and their supporting infrastructure.
- Neuro-symbolic approaches can augment and automate decision making with less risk of unintended consequences.

Current Limitations of Neurosymbolic AI

- Most neuro-symbolic AI methods and techniques are being developed in academia or industry research labs. Despite the increased availability of tools, there are still limited implementations in business or enterprise settings.
- There are no agreed-upon techniques for implementing neuro-symbolic AI, and disagreements continue between researchers and practitioners on the effectiveness of combining approaches, despite the emergence of real-world use cases.
- The commercial and investment trajectories for AI startups allocate almost all capital to deep learning approaches, leaving only those willing to bet on the future to invest in neuro-symbolic AI development.
- Currently, popular media and academic conferences do not give as much exposure to the neuro-symbolic AI movement as compared to other approaches.

Thank You!

Ask the speaker



Gartner CIO Leadership Forum

10 – 11 February 2025 | Hollywood, FL
24 – 25 February 2025 | Phoenix, AZ
10 – 12 March 2025 | London, U.K.
6 – 7 May 2025 I Sydney, Australia
9 – 10 July 2025 | Tokyo, Japan

Gartner CIO Leadership Forum is an unrivaled experience — a place where you'll access the most timely advice, actionable tools and CIO peer insights to adapt faster and stay ahead of change.

Learn more:
gartner.com/conf/cio

At this year's forum, you'll learn:

Short and long-term strategies to **lead you and your team to success** in dynamic times.

Reshape organizational structures and attitudes for greater agility.

Engage peers to co-create ways of working that remove digital friction to speed business outcomes.

Gartner®

Gartner
IT SYMPOSIUM | Xpo

Pacing Yourself in the AI Races

Two AI Races

Tech Vendor Race

Your AI Outcome Race

Your race is to deliver AI outcomes – **safely and at scale**

57% of CIOs are **tasked with leading enterprise AI/GenAI strategy and vision**

Business Outcomes

72% of IT leaders report that **employees struggle to integrate AI into daily work**

GenAI deep productivity zones

Match job complexity and job experience to unlock AI productivity

Low complexity Low experience	High complexity High experience

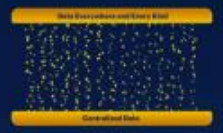
Insights

Cost estimates for GenAI can be off by 500% – 1000%

Software vendors are raising prices by up to 30% as they incorporate Gen AI

Enterprise proof of concept GenAI spend in 2023
$300K – 2.9M

Technology Outcomes



Insights

GenAI triggers a new data paradigm

Gartner Predicts

By 2026, more than 80% of software vendors will have embedded GenAI capabilities in their enterprise applications

Teenager technology

GenAI foundation models are incredibly powerful and incredibly immature

TRiSM
Trust, Risk, Security Management
Technologies

CIOs expect AI will come from **these sources:**

AI tech sandwich

How will you **customize** your tech sandwich?

Behavior Outcomes

GenAI can unleash a spectrum of emotions

Insights

Be intentional about behavioral outcomes

Agentic AI
AI agents will take action on your behalf

AI is hardening the soft sciences
Turn human insights into business value

You need human behavior experts

Get the guidance you need to set the right AI pace and safely achieve your business, technology, and behavioral goals at scale.

Download the Keynote Research Now

Gartner for CIO

How Toastmasters International Saved Over $2 Million by reviewing a proposal.

“Gartner is like a chief of staff on steroids. I get a lot of great research and information from the website, I can use them to look at contracts and recently pulled RFP templates. There's so much value in how I lead my team using Gartner, it's hard to narrow it down to one thing.”

Learn More ↗

Heidi Hollenbeck | CIO & COO

Gartner delivers actionable, objective insight to executives and their teams. Become a Client and achieve your business goals.

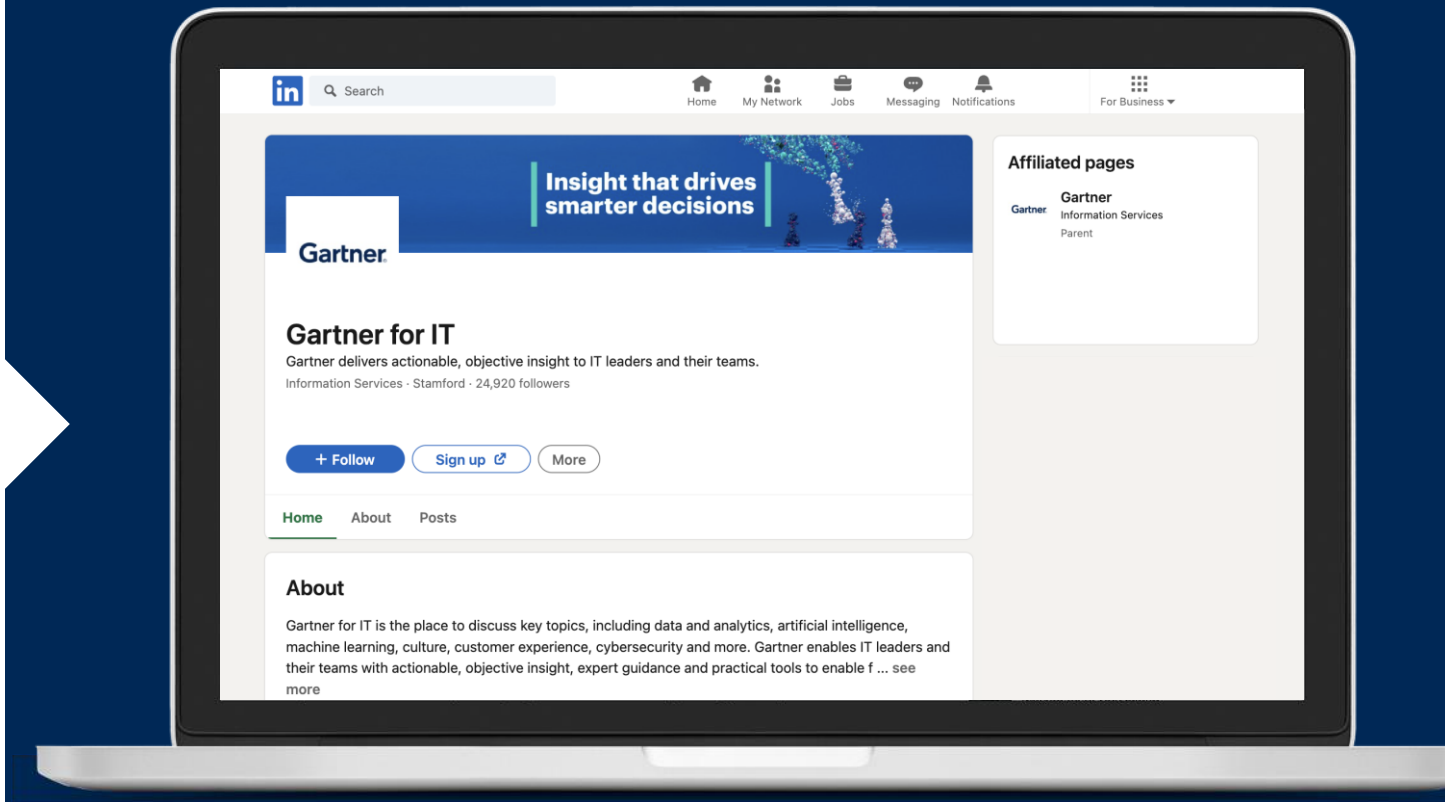
Gartner

Gartner for IT on social media

Want to stay in-the-know? Connect with us on LinkedIn and Twitter to receive the latest Gartner IT insights and updates across research, events and more. It's all curated specifically for IT leaders and decision-makers.

Follow us on

Looking for insights delivered to your inbox?
Subscribe to our bi-weekly newsletter



Gartner

Become a Client

Clients receive 24/7 access to proven management and technology research, expert advice, benchmarks, diagnostics and more.

Fill out the form to connect with a representative and learn more.

Learn More

Or give us a call: +441784614280 | +1 855 637 0291

8 a.m. – 7 p.m. ET
8 a.m. – 5 p.m. GMT
Monday through Friday

Gartner®

Get more Gartner insights

Download the research slides

View upcoming and on-demand Gartner webinars at gartner.com/webinars

Rate this session

Rate this session