

Agentic AI at Scale

Accelerating Innovation While Managing Risk



Shari Lava

Group Vice-President, AI,
Data and Automation

June 2026

Agentic AI at Scale

- 1 Where have we been?
- 2 Where are we now?
- 3 Where are we going?
- 4 How do you get there?



Where Have We Been?

A Lot Has Happened in 3 Years



A seemingly innocuous post in November 2022 is when things changed

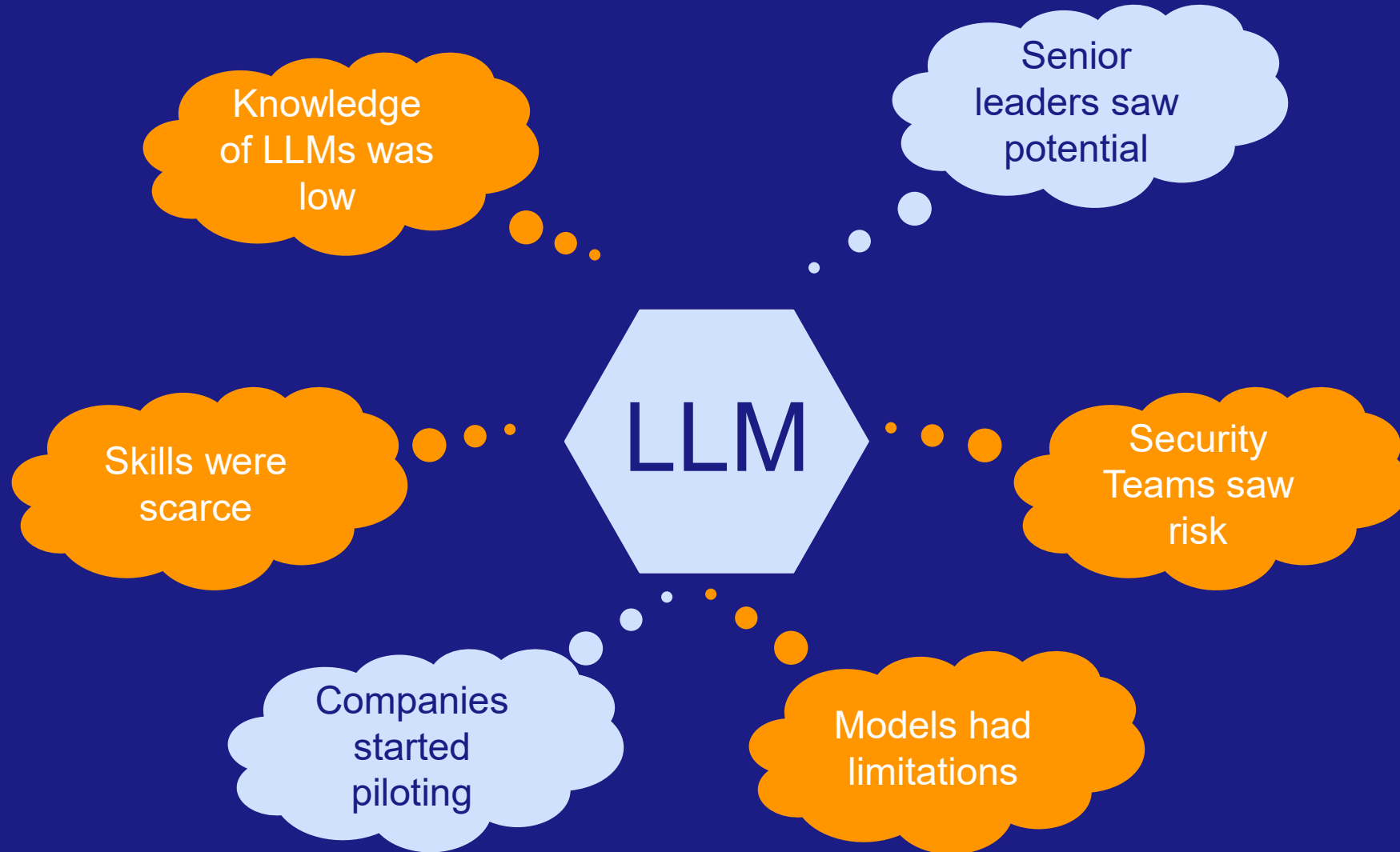


The Development of ChatGPT: A Chronology of Events

- December 11, 2015:** OpenAI is founded by Elon Musk, Sam Altman, and others with the goal of creating safe and beneficial artificial intelligence.
- June 2018:** OpenAI releases its first language model, GPT-1, with 117 million parameters.
- February 2019:** GPT-2 is released, boasting 1.5 billion parameters but limited in release due to concerns over misuse.
- June 2020:** OpenAI debuts GPT-3, which features 175 billion parameters and significant improvements.
- November 30, 2022:** OpenAI launches ChatGPT, based on GPT-3.5. The chatbot quickly reaches one million users.
- March 14, 2023:** OpenAI announces GPT-4, a multimodal model capable of processing text, images, and more.
- April 14, 2025:** GPT-4.1 is introduced with an expanded context window and various model editions.

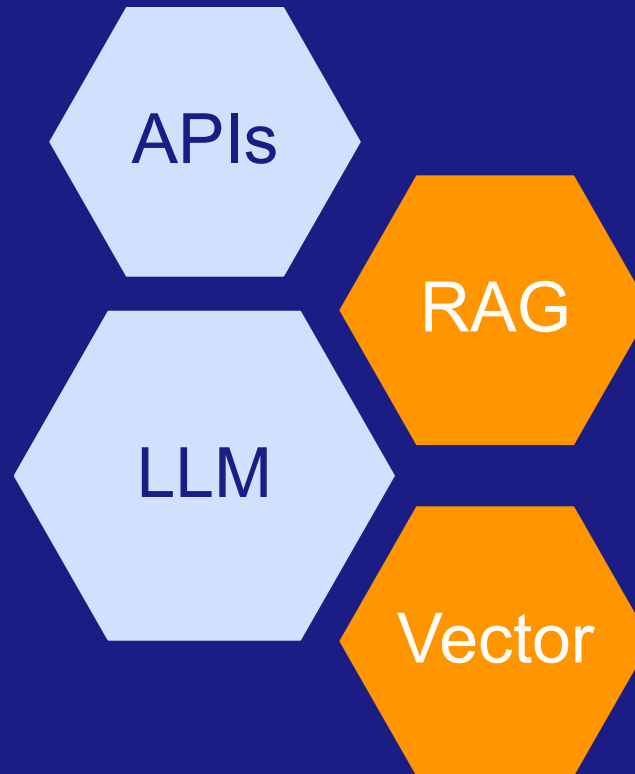


In 2023 Generative AI started to become a point of experimentation for most businesses



Teams started to make progress with making generative AI results more relevant using supporting technologies

Many vendors started to add assistive capabilities to their products



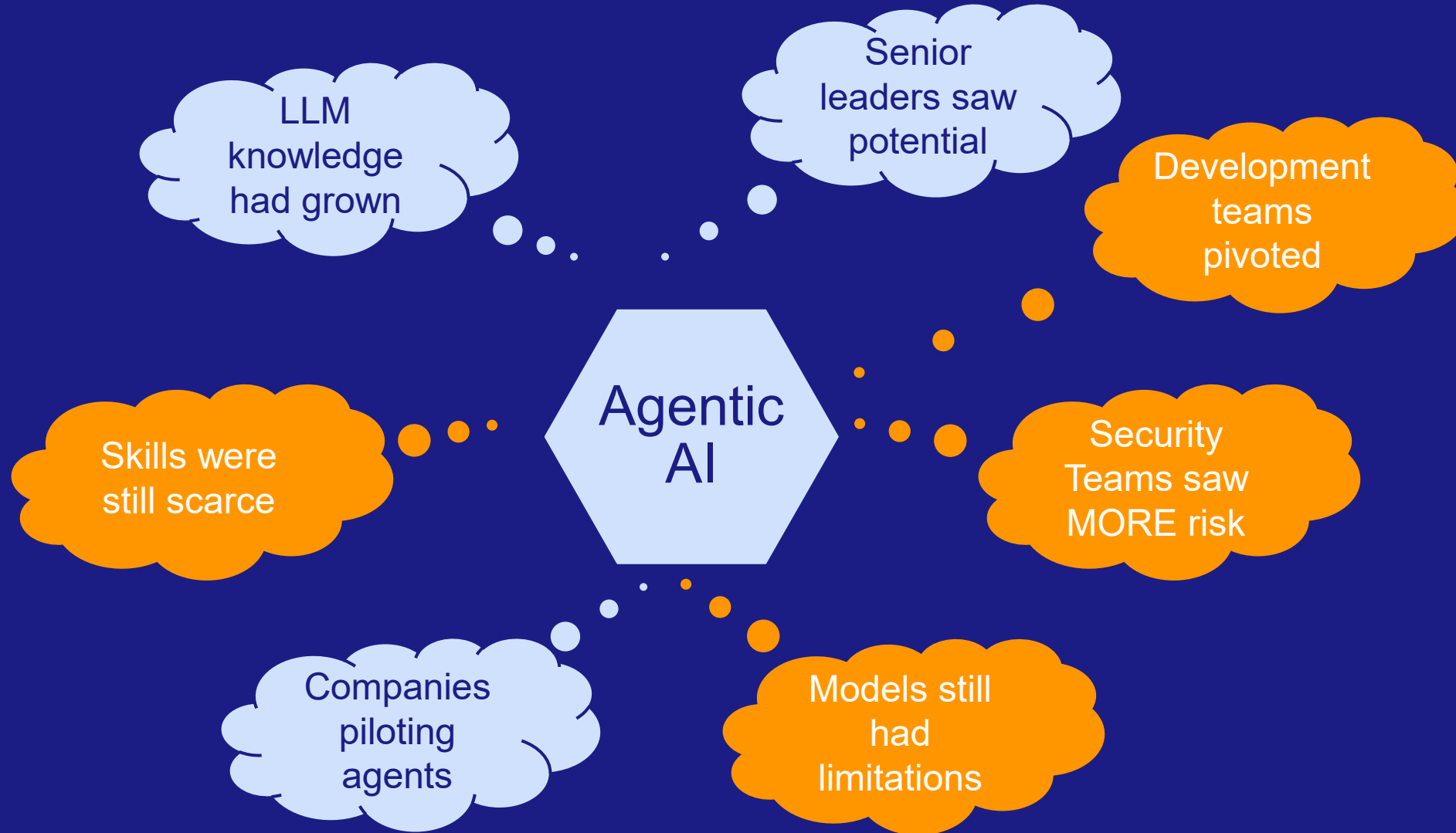
A host of new LLMs was being created

Few standards existed

Token usage a concern



By 2024 companies had experimented with GenAI, but agentic AI burst onto the scene and many companies saw it as starting over



From Assistive to Action

We are at an **inflection point** for Autonomous AI Agent Development

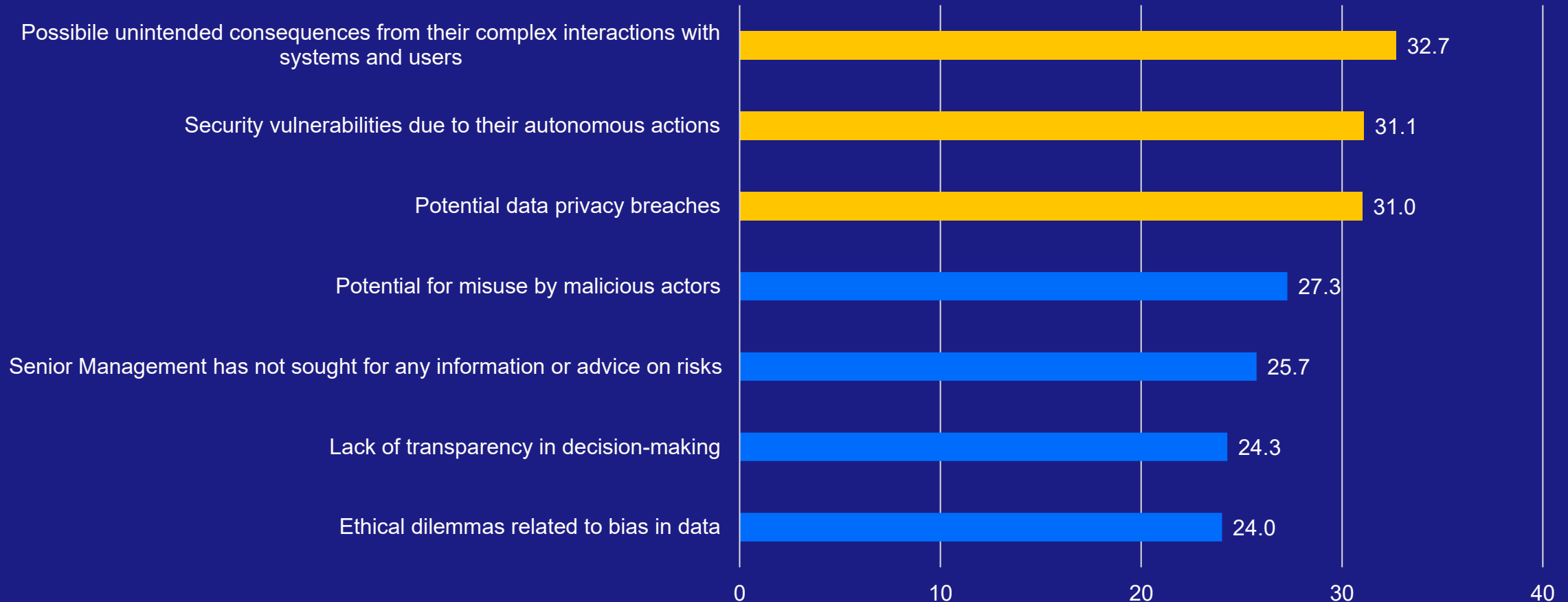


	AI Co-pilots / AI Assistants	Autonomous AI agents	AI Agent Fleets
TRUST	AI provides recommendations for decision making	Humans are "operators, supervisors and architects"	Capabilities, shares decision-making duties
TECH	Experimental agent frameworks	Productization of AgentOps tech stack	Virtual workforces of "AI" agents and marketplaces
POLICY	Responsible AI by design	Policy formalization and embedment	Productionized monitoring - "AI immune systems"



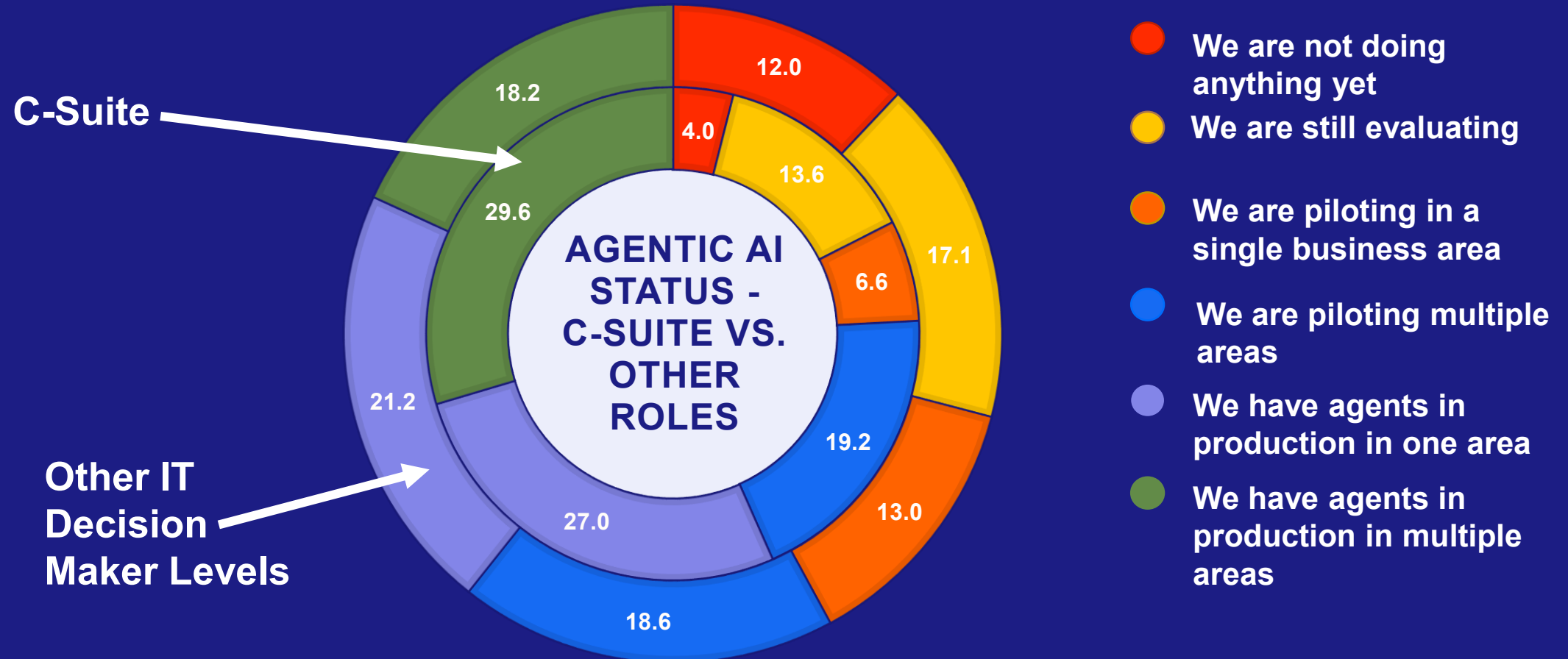
Incorporating agentic AI into the existing tech stack is a big barrier to adoption

Q. For which of the following Agentic AI-related business risks has the senior leadership most sought information from IT leadership?



There continues to be a disconnect by role, but companies are progressing

Q. Which of the following statements best describes the current state of your organization's evaluation or use of AI Agents?



Why pilots don't scale to production



Disconnected data

Gold-mine data trapped in siloed systems; pilots can't reach what they need at scale.



Fragmented AI islands

Standalone experiments and rapidly evolving models prevent holistic automation.



Undefined KPIs & ownership

Pilots launch without clear success metrics or an accountable outcome owner.



Security as an afterthought

Governance and access controls bolted on late, blocking the path to production.



Cost blind spots

No instrumentation for token spend, infra, or per-task cost; ROI can't be proven.



Expectation gap

Leadership overestimates progress while teams hit real-world deployment limits.



Where Are We Now?

Things are changing fast.



ENTERPRISES

are still evaluating
or piloting
the use of agents

47%



21%

AGENTIC AI SOLUTIONS

deployed by end of
2026

34



43

ENTERPRISES

use agents in
production across
multiple business areas

29%



50%

NOV 2025

MAR 2026



The Journey to Autonomous

We are at the **NEXT inflection point** for Autonomous AI Agent Development



	AI Co-pilots / AI Assistants	Autonomous AI agents	AI Agent Fleets
TRUST	AI provides recommendations for decision making	Humans are "operators, supervisors and architects"	Capabilities, shares decision-making duties
TECH	Experimental agent frameworks	Productization of AgentOps tech stack	Virtual workforces of "AI" agents and marketplaces
POLICY	Responsible AI by design	Policy formalization and embedment	Productionized monitoring - "AI immune systems"



Data is the next big unlock for AI

45% of AI projects to date have delivered measurable business value, but...

36.8%

are focused on improving existing AI projects, versus launching new pilots

45.9%

cite data quality and security as a top AI implementation concern

34.3%

cite integration with existing systems as a top concern

30.0%

rank data quality and availability a top **AI** investment priority for 2026

Source: Future Enterprise Resiliency & Spending Survey Wave 7, IDC, Sept 2025 (Worldwide)

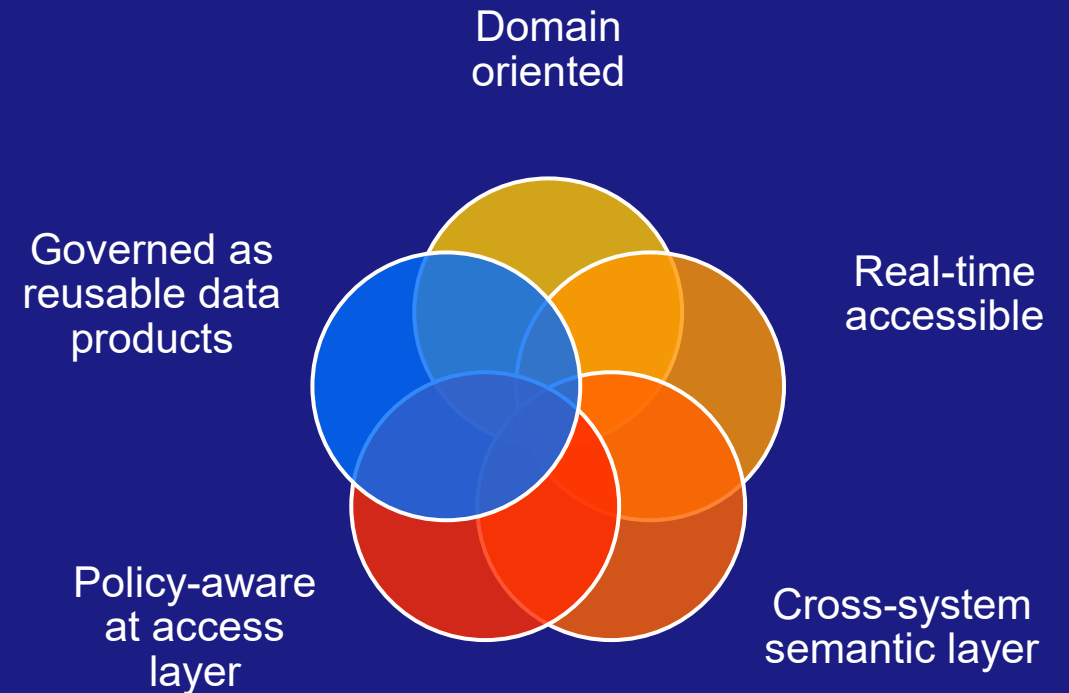


Data for applications vs agents

Application-Bound Data

- Structured around individual systems
- Transaction-optimized
- Integrated via point-to-point
- Owned by application teams
- Governed per system boundary

Agent-Ready Data



Agents don't respect application boundaries. Your data architecture still does.



Agentic AI platforms

Build:

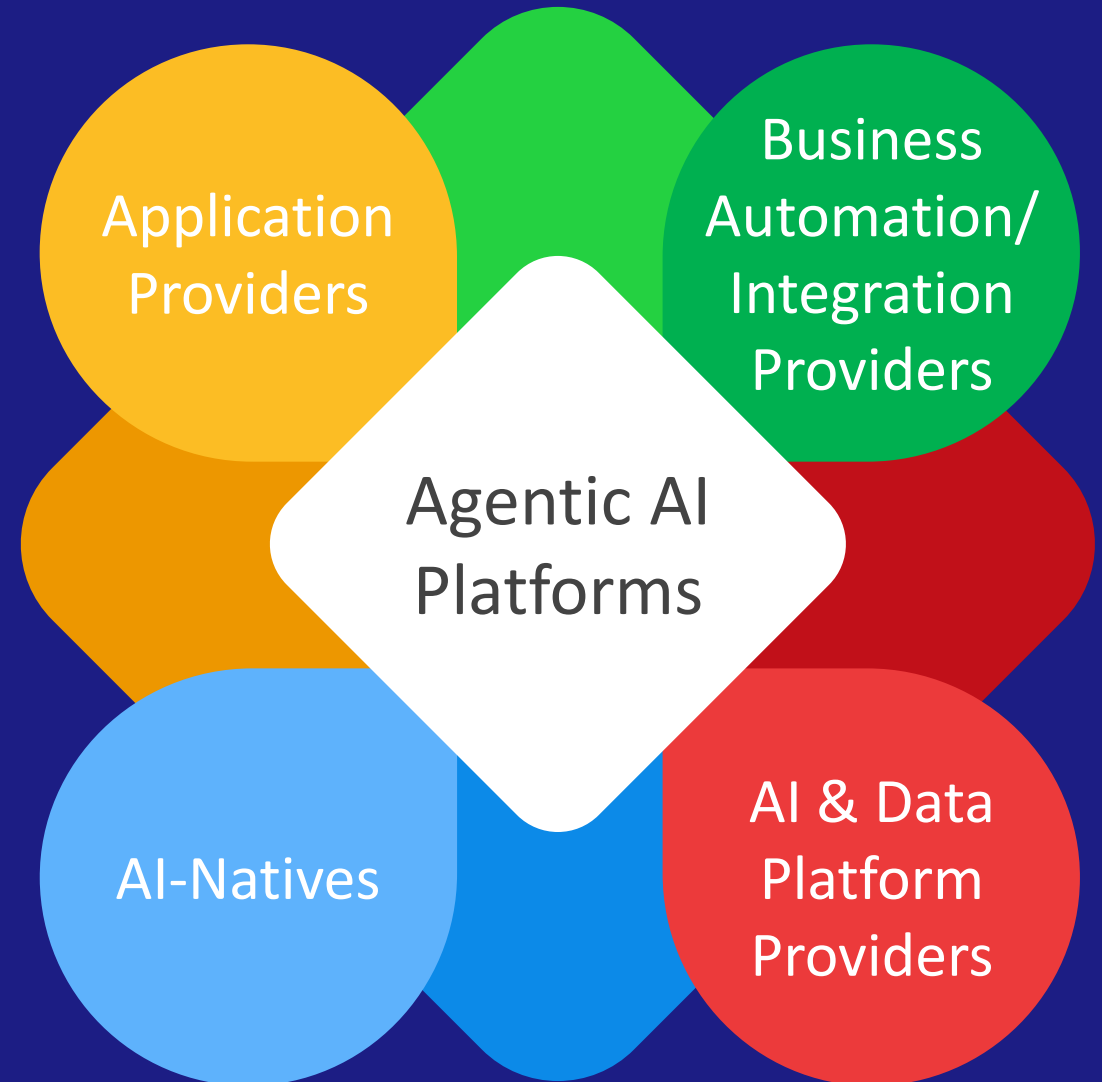
- Design & Deployment
- Integration with agent tools (API & LLM config)
- Model Routing/Orchestration
- Context engineering

Governance:

- Agentic lifecycle
- Agent Registration,
- AI Observability (inc. efficacy),
- Agent Monitoring/evals,
- Agentic identity & security,
- Agentic FinOps

Orchestration:

- Interoperability (1st party and 3rd party agents)
- Multi-agent collaboration
- Business automation (integration with business processes/workflows)

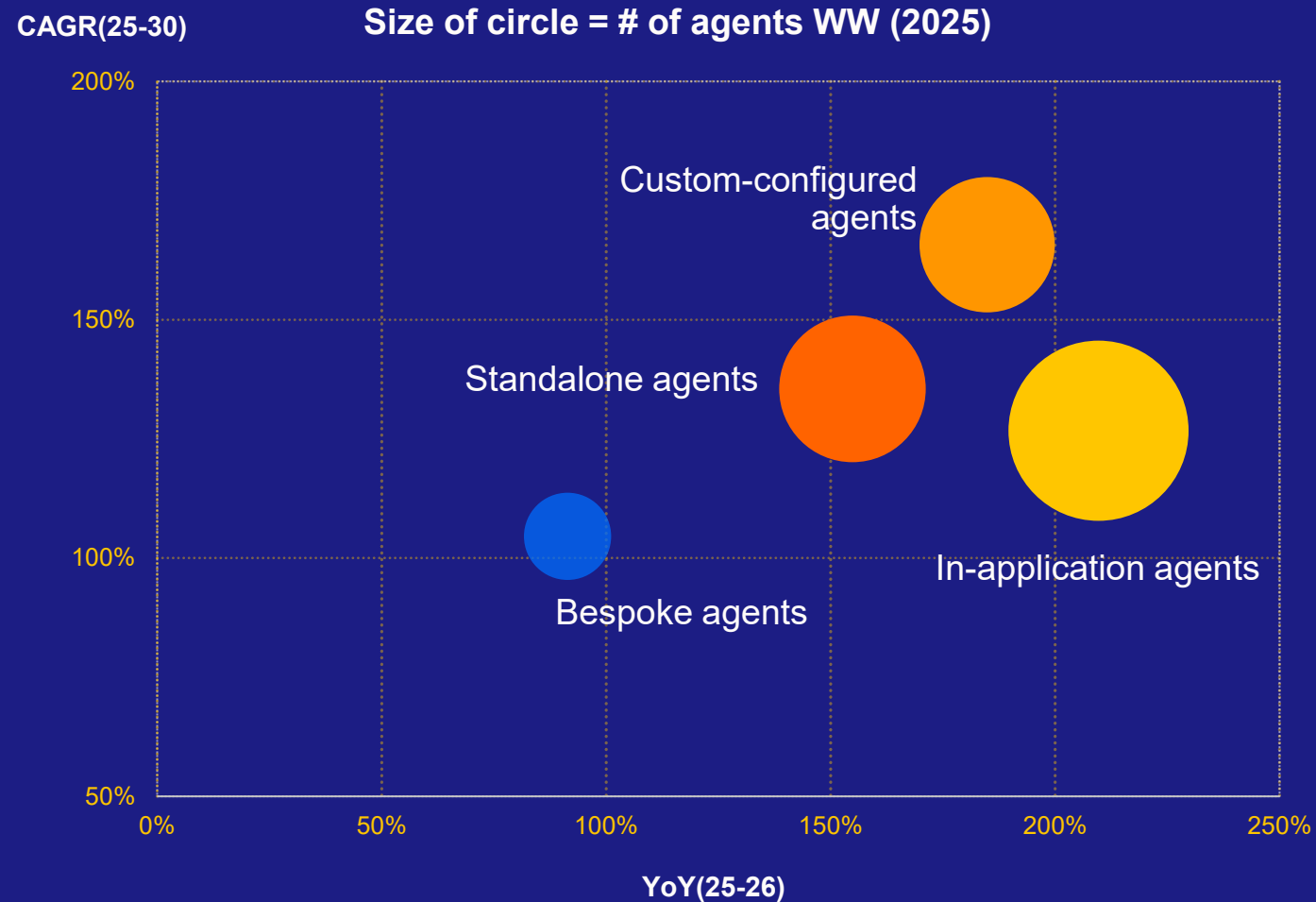


Where is this Going?

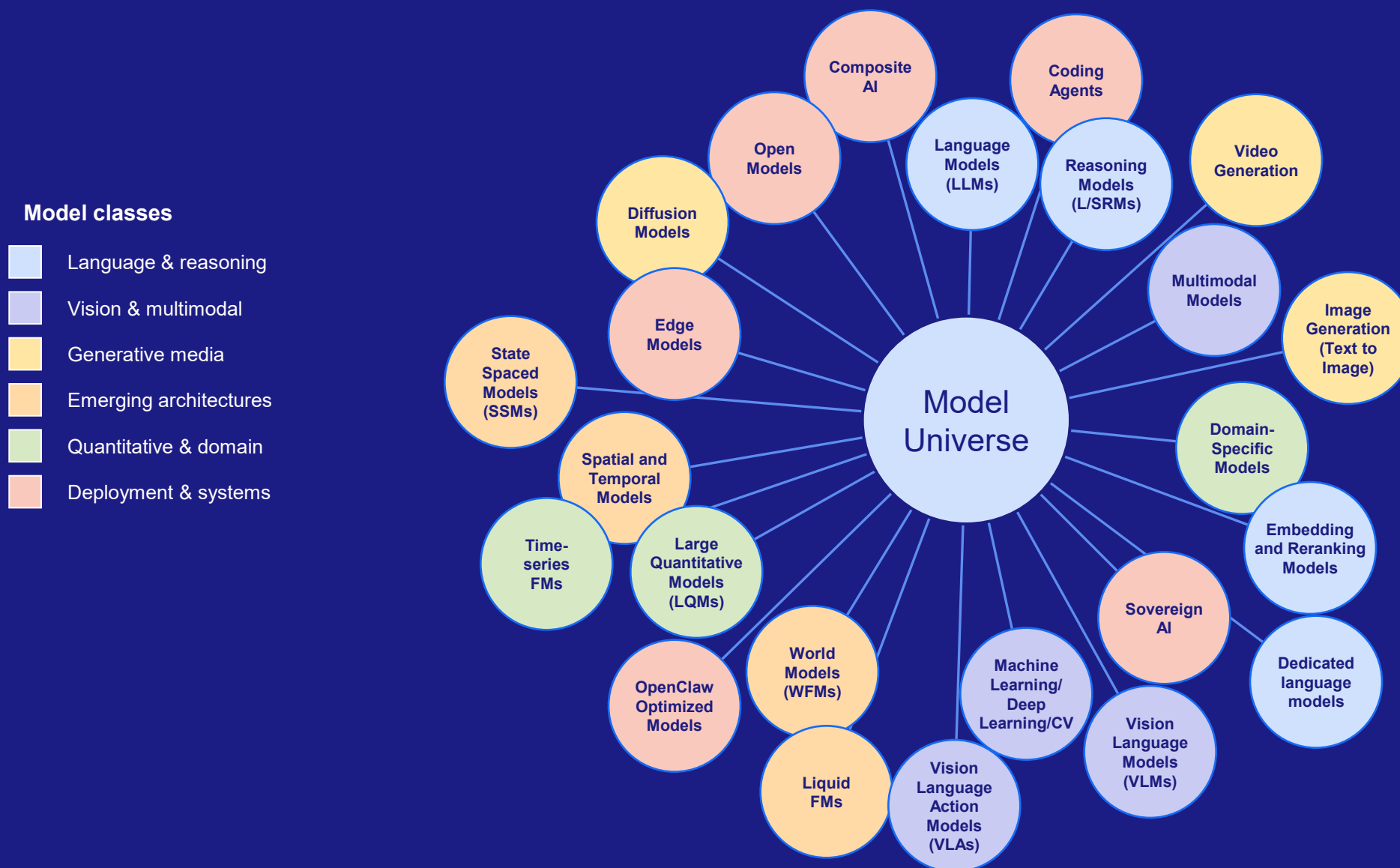
Planning for what's next is critical



1.2B Agents by 2029 for 217B daily actions



The Universe (of Models) is Expanding



The Quest For Autonomous

We are at an **inflection point** for Autonomous AI Agent Development



	AI Co-pilots / AI Assistants	Autonomous AI agents	AI Agent Fleets
TRUST	AI provides recommendations for decision making	Humans are "operators, supervisors and architects"	Capabilities, shares decision-making duties
TECH	Experimental agent frameworks	Productization of AgentOps tech stack	Virtual workforces of "AI" agents and marketplaces
POLICY	Responsible AI by design	Policy formalization and embedment	Productionized monitoring - "AI immune systems"

Source: Tech Buyer Presentation - Document number:# US52518424, Aug 2024; Adapted from Aura Ventures

Distribution is forbidden without written consent of IDC. All rights reserved.



Agentic AI platforms

Build:

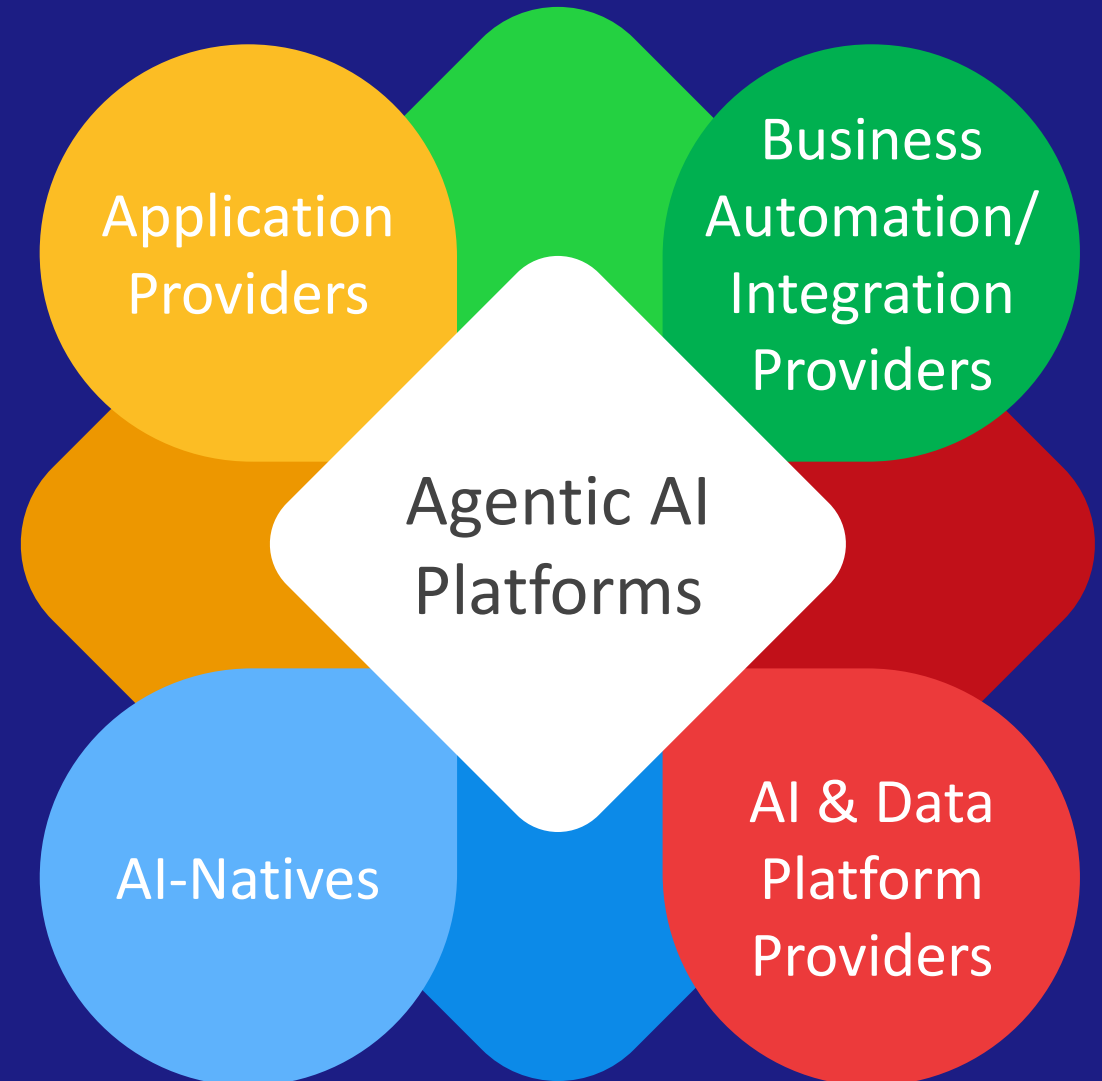
- Design & Deployment
- Integration with agent tools (API & LLM config)
- Model Routing/Orchestration
- Context engineering

Governance:

- Agentic lifecycle
- Agent Registration,
- AI Observability (inc. efficacy),
- Agent Monitoring/evals,
- Agentic identity & security,
- Agentic FinOps

Orchestration:

- Interoperability (1st party and 3rd party agents)
- Multi-agent collaboration
- Business automation (integration with business processes/workflows)



“

Agentifying a process where rules-based automation works well is like putting a Ferrari engine in a lawnmower

~Gemini 3.1, May 2026



Order to cash — agentic architecture



How do you get there?

This is not a single team or project



IDC MaturityScape: AI-Fueled Organization 2.0

FIVE MATURITY STAGES



FOUR MATURITY DIMENSIONS, 15 SUB-DIMENSIONS



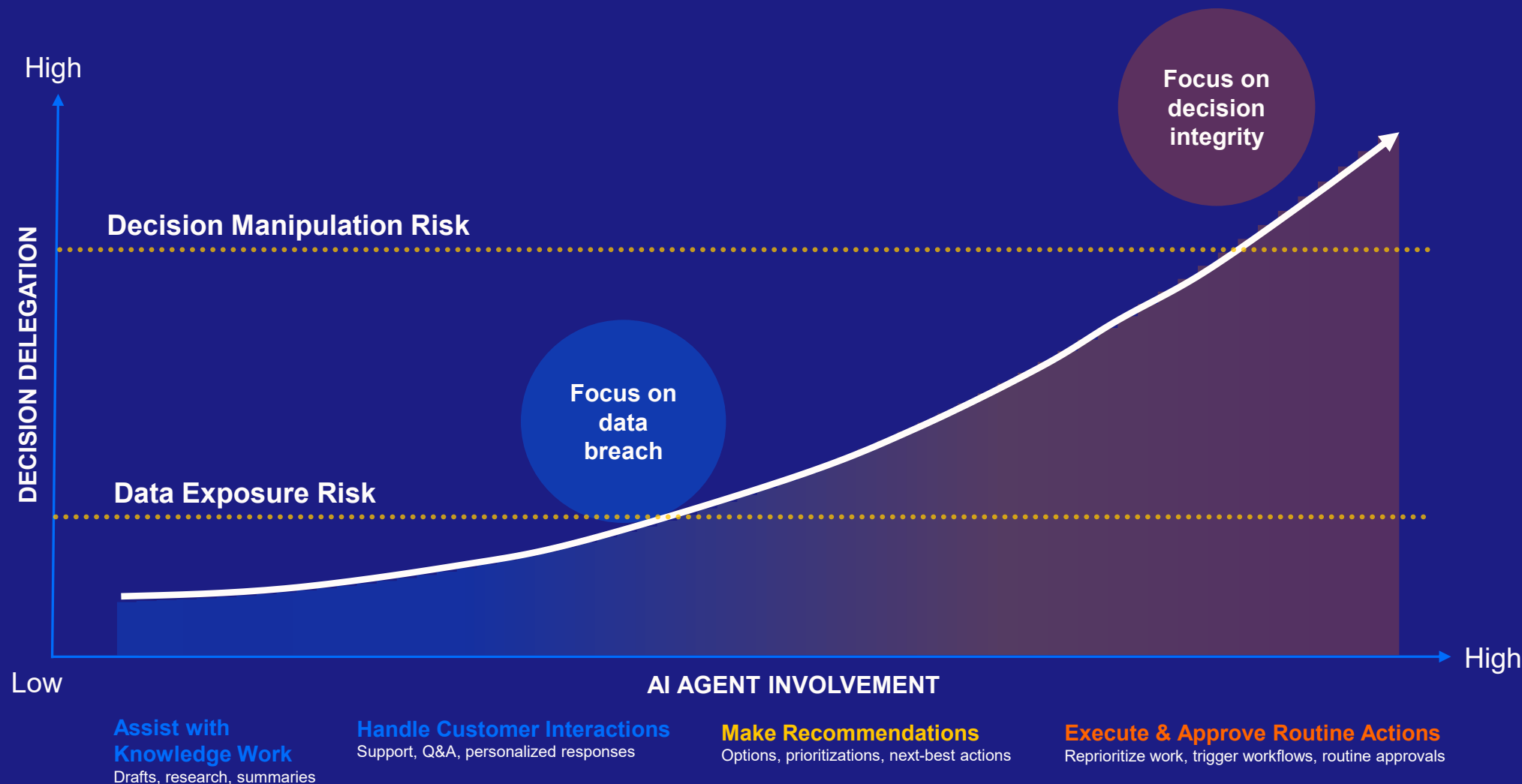
Source: IDC MaturityScape: AI-Fueled Organization 2.0, April 2026



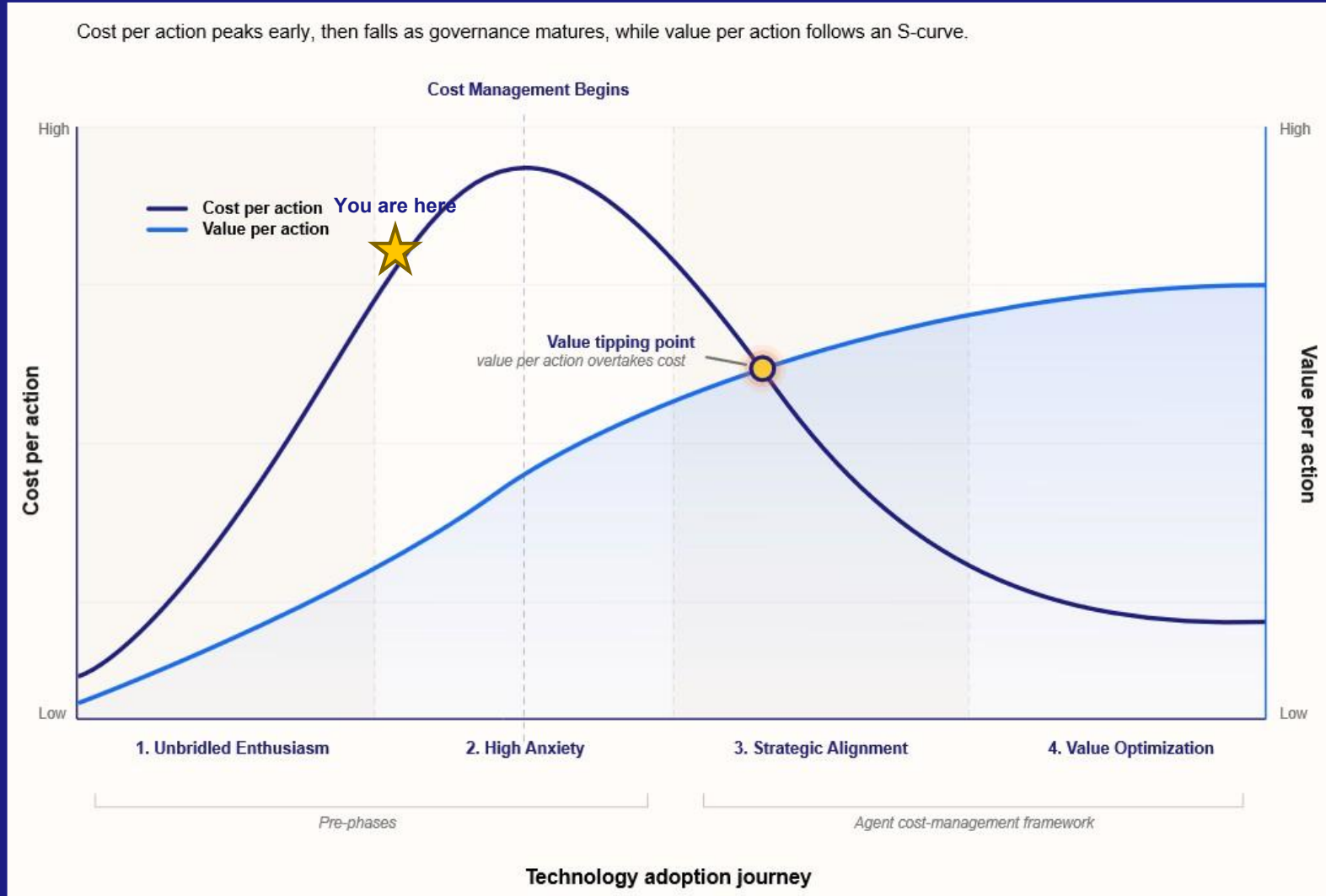
1.2 billion AI agents worldwide by 2029



Agentic AI security isn't just about data



Some early adopters are starting feel the pain of unmanaged AI Cost



Top 3 Barriers to Evaluating Agentic Pricing:

- *Lack of Clarity on Pricing*
- *Difficulty Budgeting for Tokens*
- *Unpredictable Usage Costs*

AI cost management will be a new discipline that requires ongoing, strategic planning and management

Source: IDC, FERS Survey 2026 Wave 3, May 2026

Source: IDC, Effective Agent Cost Management: A Practitioner's Framework for Planning, Sourcing, Implementation, and Operations, June 2026 (Expected)



Excerpt of AI cost management advice: Match models to tasks

Task Type	Appropriate Model Tier	Cost Implication
Simple classification, extraction, routing	Small / fast model	Materially lower cost than frontier
Summarization, reformatting, translation	Small-to-mid model	Moderately lower cost than frontier
Reasoning, planning, complex generation	Mid or frontier model	Baseline cost
Long-context analysis, advanced coding	Frontier (large context)	Premium cost; optimize carefully
Real-time user-facing responses	Fast model (latency matters)	Prioritize speed over cost efficiency

Right-size the model to the task - it's the single biggest lever on inference cost.



Key Takeaways for Technology Leaders

1



Shift from scattered pilots to one value-led road map

Fewer than half of AI projects deliver measurable value. Isolated pilots automate point tasks, not whole business processes. Prioritize a road map that transforms end-to-end workflows.

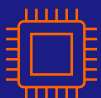
2



Make governance a core pillar before you scale agents

AI agents in production are climbing from ~40 to 50 per organization. Put responsible AI, risk, and data governance in place now. Ensure it is a coordinated effort.

3



Right-size models and modernize your data foundation

Match each task to the right model tier to cut inference cost, and rebuild data architecture to be agent-ready, not application-bound.



