Corporate Buyers' Guide to Enterprise Intelligence Applications - Q4 2024

Navigating the landscape of GenAI/LLM trends, use cases and vendors

GAI Insights

Sign up for report updates

About GAI Insights - research offerings

Product Offerings:

- **AI Strategy Development**
- **GenAI Training**
- **Benchmarking**
- **AI News Monitoring**
- **Executive Briefings**
- **Annual Conference**

Demonstrated leadership in GenAI (GAI Insights' milestones)

➤ The first and only industry analyst firm exclusively dedicated to GenAI, with a team of 8 AI analysts

➤ The first (and only) firm whose two principals were appointed as executive fellows at Harvard Business School, specifically focusing on GenAI

➤ The first firm to host a conference that solely centers on Enterprise GenAI

➤ The first firm to feature AI as a co-host of its conference

➤ The first firm publish Corporate Buyers Guide to LLM/GenAI

Paul Baier

CEO and Co-founder

GAI Insights

25 year software entrepreneur including 5 years at an AI company
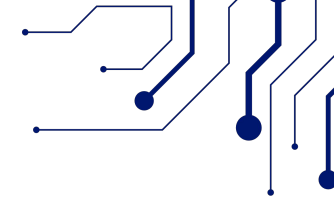
Earned MBA at Harvard

Dr. John J. Sviokla

Principal and Co-founder

GAI Insights

Former PwC Partner (10 yrs) and

HBS professor (12 years)

Wrote 3 books, over 90 HBR articles

Served on multiple Boards

Earned DBA, Master and BA from Harvard

GAI Insights

This report offers unique value

Research methodology

➤ **In-depth interviews** with 25 AI executives from our leadership forum, conference speakers, and survey participants

➤ **Product evaluations** through vendor briefings

➤ **Comprehensive analysis** of enterprise use cases, with detailed LLM implementation data

➤ **Expert perspectives** from 25 consultants and industry analysts

➤ **Proprietary vendor rating framework** assessing market traction, execution capabilities, and organizational strength

➤ **Key trends** identified through our daily AI news coverage and work with companies

Value for each audience

➤ **Enterprise leaders**
  • Validate AI investments with real-world use cases
  • Fast-track vendor selection
  • Build your multi-year roadmap confidently

➤ **Investors**
  • Track market opportunities and spending trends
  • Identify high-growth AI segments
  • Evaluate competitive landscape

➤ **AI vendors**
  • Sharpen market positioning
  • Align with enterprise buying patterns
  • Optimize sales strategies

GAI Insights

Table of contents
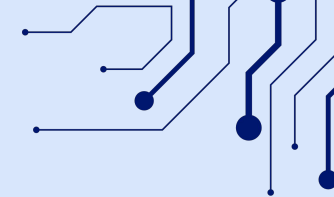
Background and overview

Executive summary

Section 1: Adoption and Use Case Trends

- Growth and adoption of GenAI
- WINS framework
- GenAI use cases adopted by prominent enterprises
- GenAI as an "Intelligence leverager"
- Eight steps to GenAI top quartile performance

Section 2: Vendor and technology trends

- Investment in GenAI and training costs
- Vendor trends and roadmap
- Market opportunity and outlook

Section 3: Enterprise Intelligence Applications (EIA) emerging leaders

Section 4: 2025 purchase considerations

Section 5: Vendor profiles

Section 6: Appendix

GAI Insights

Published Q4, 2024

GenAI vendor landscape consists of multiple vendors that play varying roles

Foundation model vendors	Cloud vendors	No-code & other deployment solutions	Chip + LLM solutions
AI21 labs	aws	arcee.ai	NVIDIA
ANTHROP\C	databricks	CustomGPT.ai	SambaNova SYSTEMS
cohere	Google	deepset	
Inflection AI	IBM	glean	
Liquid	Microsoft	H2O.ai	
Meta	ORACLE	LangChain	
MISTRAL AI_	snowflake	LlamaIndex	
OpenAI		PRYON	
		QuantumGears LLM-NATIVE INTEGRATION	
		seekr	
		together.ai	
		vectara	
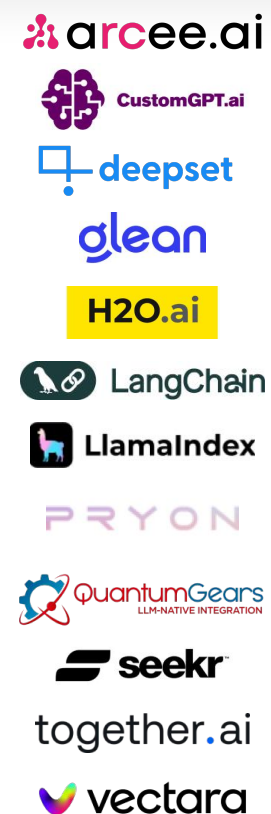
GAI Insights

Executive Summary

GAI Insights

Q4 2024 Emerging Leaders in Enterprise Intelligence Applications

Scatter plot with axes "Market Readiness" (Low to High, horizontal) and "Strategic Value" (Low to High, vertical). Positioned logos include Meta, OpenAI, ANTHROP\C, Microsoft, cohere, amazon, and CustomGPT.ai, along with numerous "Vendor" labels.
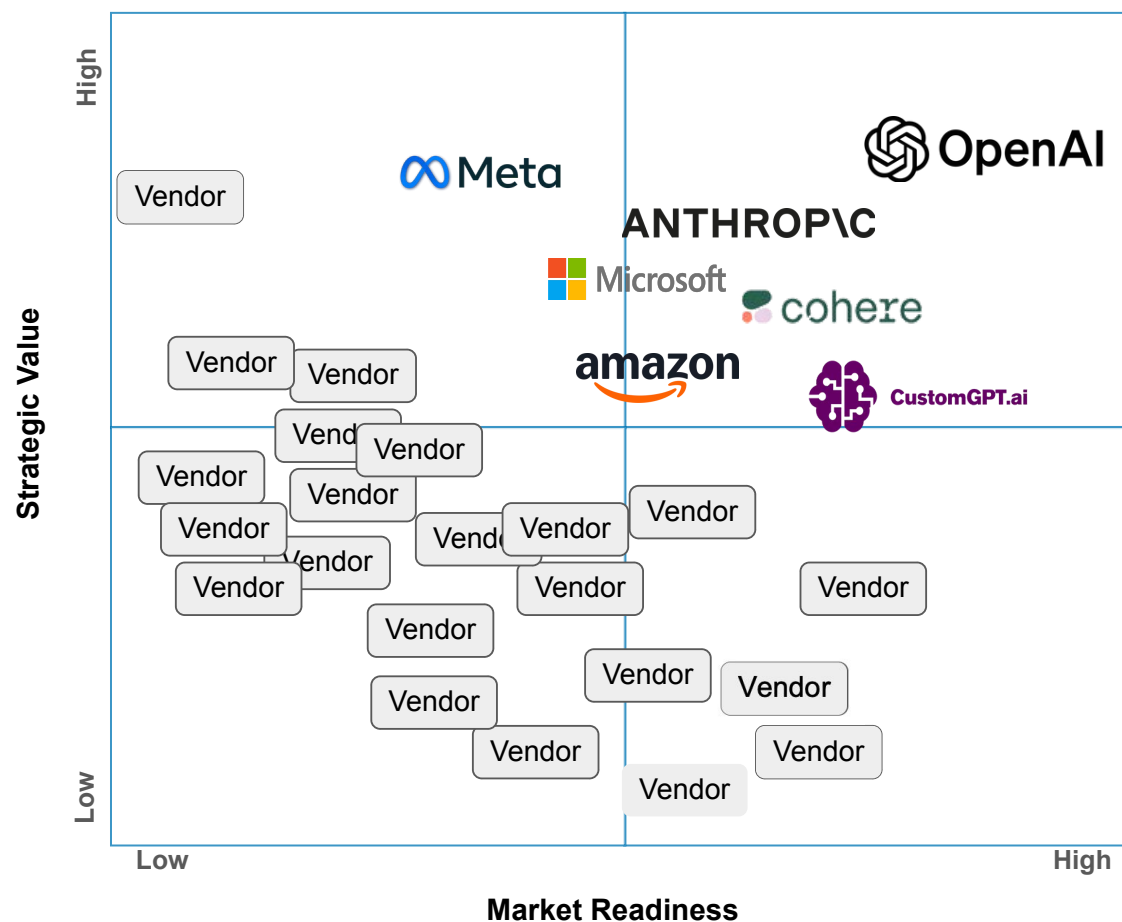
Content highlights

- ➤ **Review of 29 vendors:** Focused on the most common enterprise GenAI use cases, including employee chatbots, customer service, sales support, marketing, and supply chain optimization

- ➤ **Use case-oriented analysis:** Assessed through a use case perspective rather than solely a technology approach (e.g., foundation models or vector databases)

- ➤ **Hyperscaler products:** Evaluated proprietary products from Google, Microsoft, and Amazon (e.g., Microsoft AI Studio, AWS Titan), not products they distribute

- ➤ **Extensive research:** Findings based on 100 interviews and 500 hours of research

- ➤ **Vendor-neutral approach:** This is an independent analysis with no "pay-for-play" influence

- ➤ **Proprietary scoring system:** Scoring weighted heavily by customer traction and company capabilities

GAI Insights

This analysis is based on the following products

Vendor	Product	Vendor	Product	Vendor	Product	Vendor	Product
AI21 Labs	Jamba – Large, Mini (P)	Deepset	deepset Cloud, deepset Studio Haystack (O)	Liquid AI	LFM 1.3B (P) LFM 3.1B (P) LFM 40B MoE (P)	Oracle	OCI GenAI Service
Anthropic	Claude: Sonnet, Haiku, Opus (P)	Glean	Glean	Llama Index	Llama Index	Pryon	Pryon RAG Suite Pryon Ingestion Engine Pryon Retrieval Engine Generative Engine
Arcee AI	Orchestra Small models family (O) SuperNova 72B (P)	Google	Gemini 1.5 Pro, Flash (P) Gemma 2 9B – (P)	Meta	Llama 3.1 8B, 70B, 405B (O)	Quantum Gears	QS SecureGPT QS Contracts QS Benefits
Amazon	Titan (P) Bedrock	H2O.ai	H2O AI Cloud H2O Driverless AI Open-source libraries	Microsoft	Azure AI Studio Azure OpenAI Service Phi-3 (O)	SambaNova Systems	DataScale SambaNova Suite SambaNova Cloud
Cohere	R (P)	IBM	Granite(P) Watsonx	Mistral AI	Ministral 3B, 8B (O) Mistral Large 2, Small (P) Pixtral, Pixtral Large, Mathstral Codestral Mamba. Le Chat	Seekr	SeekrAlign SeekrFlow
Custom GPT	CustomGPT	Inflection AI	2.5 (P) 3 Enterprise (P)	Nvidia	Nvidia AI Enterprise suite Nvidia AI Foundation Models (P) Nvidia Omniverse platform Nvidia NeMo	Snowflake	Arctic Snowflake Cortex Snowflake ML Pre-built AI functions Snowpark Container Services Snowflake Marketplace
Databricks	DBRX, MosaicAI (P)	LangChain	Lang Chain Lang Graph Lang Smith	OpenAI	GPT-4 and -o1 (P)	Together.ai	Decentralized Cloud Together Inference
						Vectara	Vectara Platform Mockingbird LLC Neural Rank

OEM Developer Types   (P) Proprietary Model   (O) Open-Source Model

Executive summary – enterprise lens
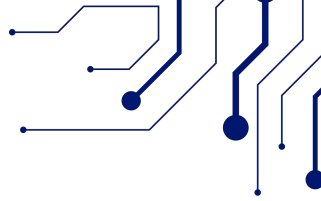
Early market with a focus on educate

- The GenAI market remains in its early stages, with only 5% of applications currently in production
- Most organizations remain in the "Educate" phase of the Educate, Apply, Transform (EAT) adoption framework

Highest RoI for firms "in the crucible"

- Firms "In the crucible" of the WINS model achieve the highest ROI from adoption
- Coding, customer service, and contract management deliver the strongest ROI, often using a Retrieval-Augmented Generation (RAG) approach

Own your own intelligence

- There is growing interest in own your own intelligence (OYOI) and on-premise solutions due to the need for greater data privacy, regulatory compliance, and control over proprietary data.
- These solutions reduce cloud dependency, lower costs, and enable tailored AI development for competitive advantage.

GenAI remains a priority

- Many companies select GenAI solutions from their existing technology vendors, particularly cloud providers
- A pragmatic approach prevails regarding "AI agents" as they are not a top priority despite the surrounding hype
- GenAI remains the leading budget priority for CIOs and CTOs

GAI Insights

Executive summary – market outlook
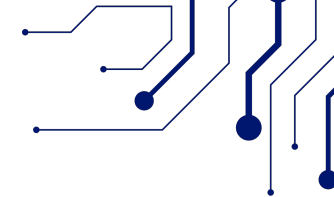
2025 is expected to be the year of "GenAI in production," following the pilot phases of 2023–2024.

AI agents continue to see very slow adoption

On-premise adoption is increasing steadily

Vendor messaging is improving as the industry grows weary of inflated claims.

LLM deployment is expected to become more common on mobile devices, edge devices, laptops, and on-premise systems.
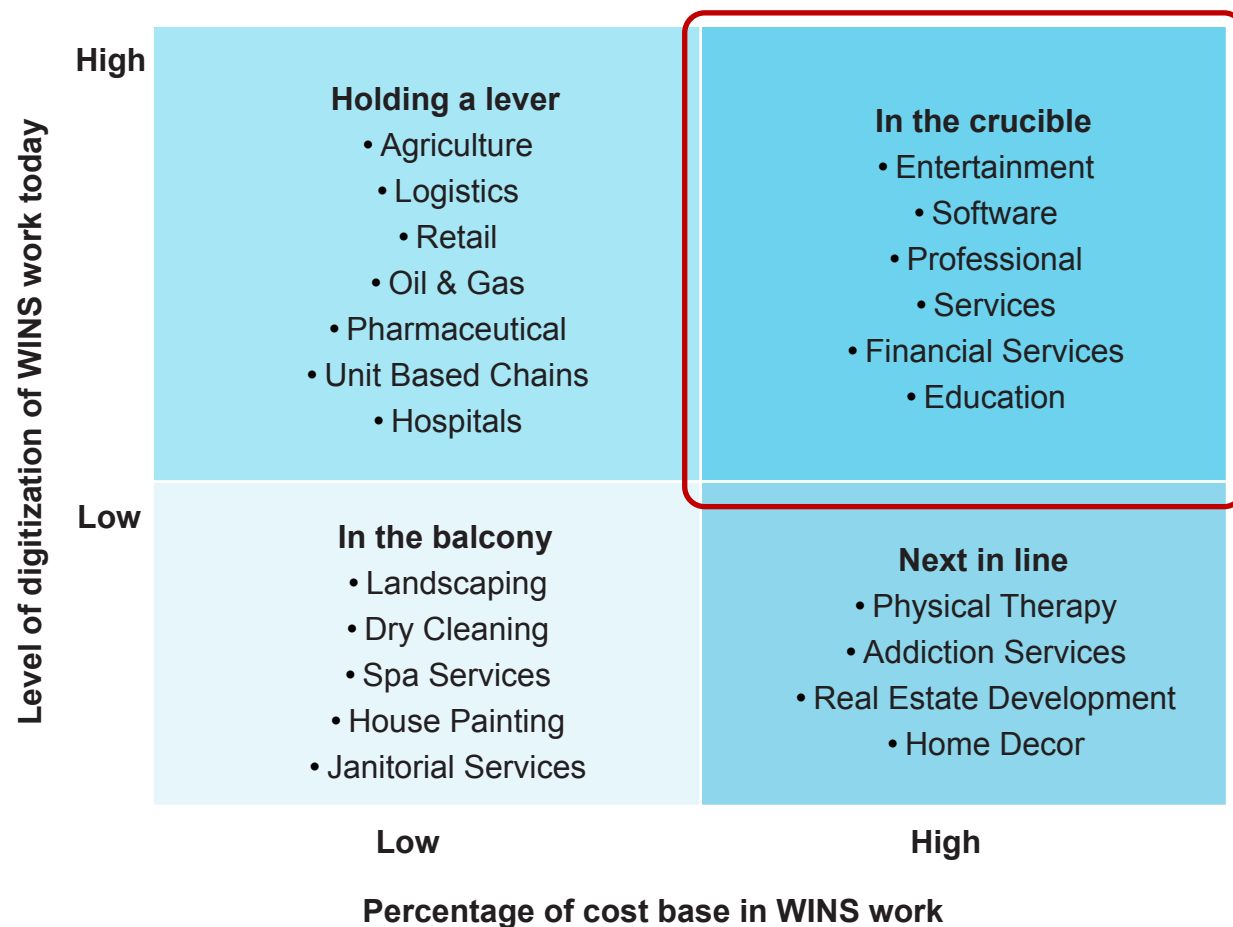
GAI Insights
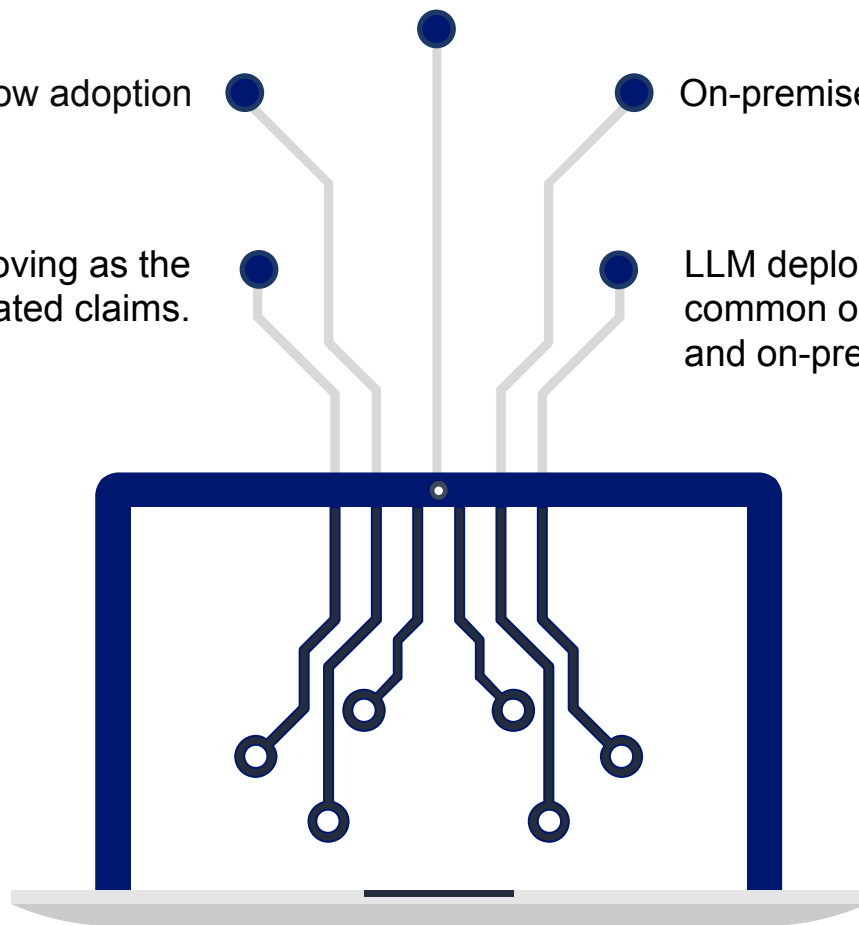
Section 1:

Adoption and Use Case Trends

GAI Insights

Adoption is highest by firms "in the crucible" in our WINS framework

The WINS "framework with industries"

Level of digitization of WINS work today

High

Holding a lever
- Agriculture
- Logistics
- Retail
- Oil & Gas
- Pharmaceutical
- Unit Based Chains
- Hospitals

In the crucible
- Entertainment
- Software
- Professional
- Services
- Financial Services
- Education

Low

In the balcony
- Landscaping
- Dry Cleaning
- Spa Services
- House Painting
- Janitorial Services

Next in line
- Physical Therapy
- Addiction Services
- Real Estate Development
- Home Decor

Low   **High**

Percentage of cost base in WINS work

- ➤ WINS stands for "words, images, number, and sounds" and is a subset of knowledge work
- ➤ "In the crucible" industries and firms feel the biggest revenue and profit impact in next 2 years
- ➤ These firms are aggressively investing in GenAI. Examples include

GAI Insights

Customer support, coding, summarization, and text generation are top use cases of GenAI being pursued by top enterprises

| Enterprises	GenAI vendor used	Use case
Klarna	OpenAI	Customer service chatbot and HR software
Morgan Stanley	OpenAI	Information retrieval for wealth management managers
Pfizer	ANTHROP\C aws	Search documents by voice command and chatbot
servicenow.	Microsoft NVIDIA OpenAI	Generating sales emails and code generation
AT&T	Microsoft	Customer Service chatbots and content generation for marketing materials
Uber	Google OpenAI	Demand forecasting and dynamic pricing optimization

GAI Insights

Section 2:
Vendor and Technology Trends

AI startups captured nearly a third of all venture funding in Q3'24
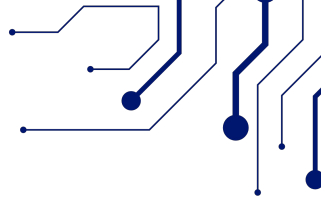
24% **QoQ**

Increase in AI deal count

$23.5M

Average AI deal size in 2024

13

New AI unicorns

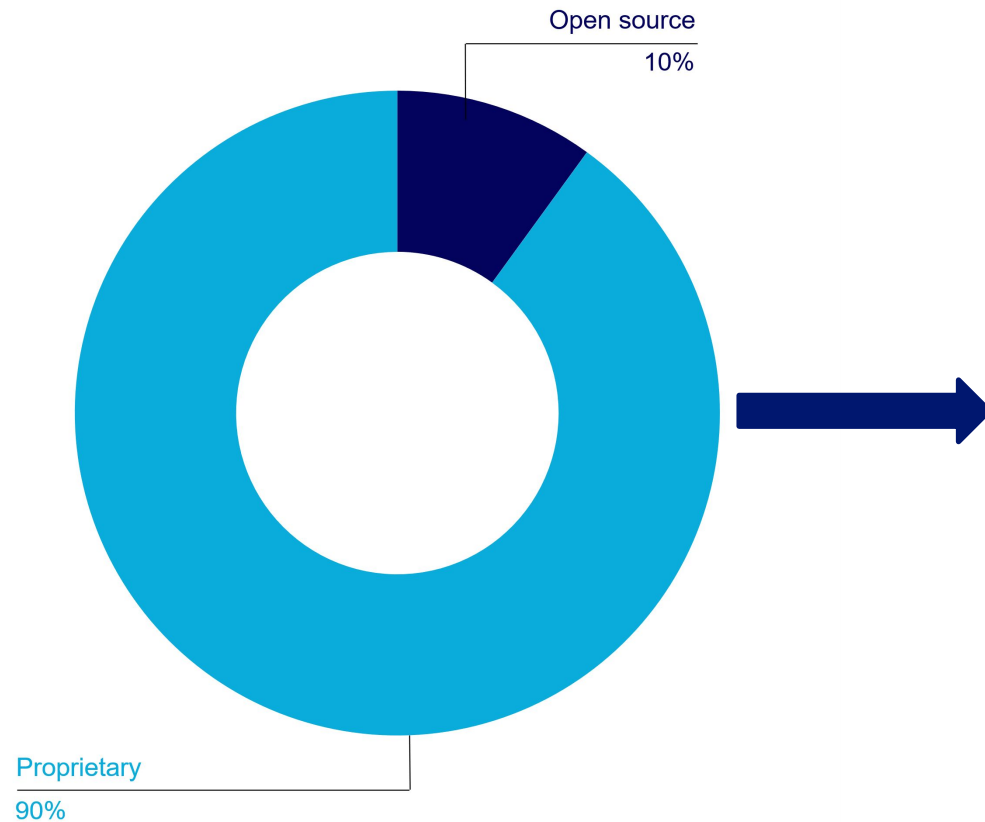**48%**

Decrease in AI M&A exits

$11.4B

US AI funding

GAI Insights

Proprietary models and OpenAI dominate market share

Estimated LLM market share

Open source
10%

Proprietary
90%

Proprietary LLMs market share

Open AI / Azure	65%
AWS/ Anthropic	11%
Cohere	8%
Google/GCP	8%
A121 Labs	1%
Nvidia	1%
Databricks	1%
Others (IBM, Inflection AI,etc.)	5%

Ir

0%   20%   40%   60%
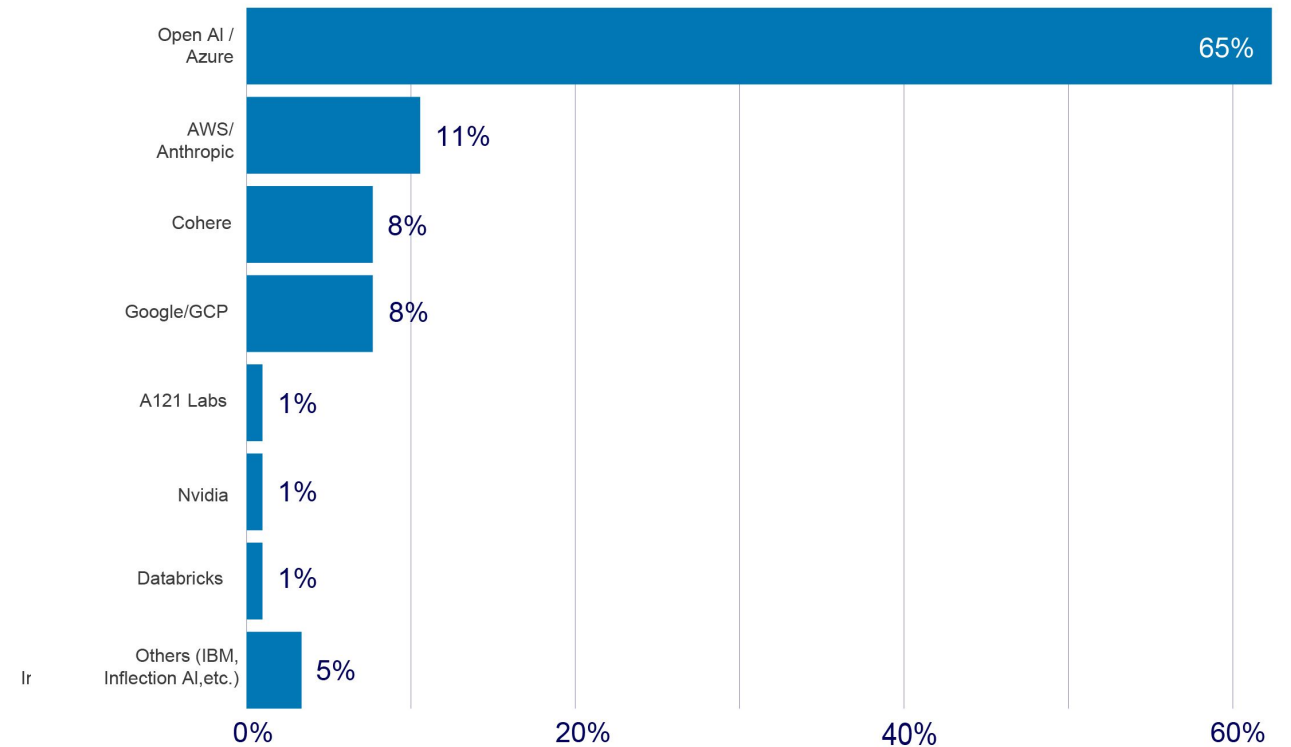
GAI Insights

Section 3:

Enterprise Intelligence Applications (EIA) Emerging Leaders

GAI Insights

2024 Emerging Leaders in Enterprise Intelligence Applications

Enterprise Intelligence Applications (LLMs, etc.)
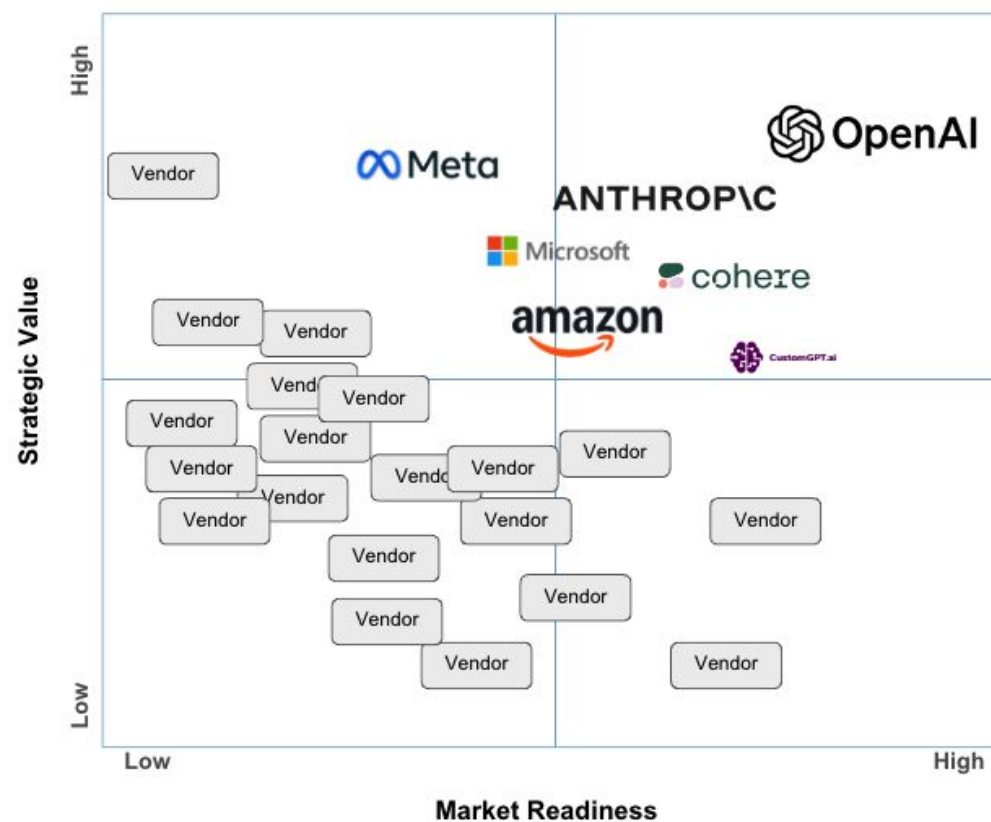
Section 5:

Vendor Profiles

GAI Insights

Overview   AI21 Labs, founded in 2017, focuses on developing advanced large language models (LLMs) specifically tailored for enterprise applications. With around 290 employees and headquarters in Tel Aviv, Israel, they provide both general and task-specific capabilities. Their latest family of models, Jamba, utilizes innovative architectures like the hybrid SSM-Transformer.

Leadership   Amnon Shashua, Ori Goshen, Yoav Shoham

Key investors	Insight Partners, Accel, General Catalyst, Google, Nvidia
Total funds raised	$326M

Problem they are trying to solve: AI21 is addressing the challenge of turning LLMs into scalable, production-grade solutions for enterprise use. The main issues they solve are related to deploying LLMs in business environments, where cost, scalability, and performance with long documents are critical concerns. They aim to overcome the inefficiencies of traditional transformer-based models, particularly in handling long-context tasks, reducing latency, and minimizing resource consumption.

GAI Insights view

- **Long Context Handling**: Jamba supports an effective context window of up to 256k tokens, allowing enterprises to process long documents or multiple documents for complex comparisons and integrations.

- **Unique architecture:** Jamba's architecture is a combination of transformers and Mamba, designed for speed and efficiency. This hybrid model improves memory efficiency, enabling it to run on smaller machines and reduce costs

- **On-Prem and Air-Gapped Deployments**: Jamba can be deployed on-prem or in air-gapped environments, a feature many enterprises value for data security and control.

- **Distribution and market share**: AI21 Labs has developed robust foundation models and established an initial customer base. While it has made significant strides, it continues to expand its distribution and market presence.
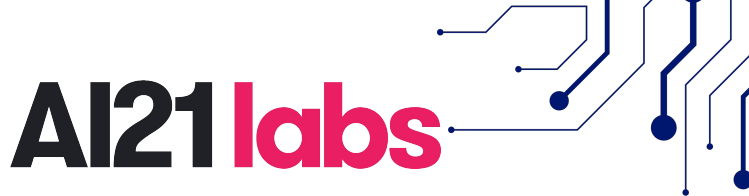
Public customers

Product name	Description	Pricing
Jamba 1.5 Mini	The Jamba 1.5 Mini features 12 billion active parameters and 52 billion total parameters, with a 256K context window. Built on the Jamba architecture, it merges Mamba and Transformer layers to deliver both high performance and efficiency, particularly for tasks requiring long-context processing. This open model family makes powerful, low-latency, high-quality LLMs accessible, while excelling at handling extended contexts.	**Input tokens**: $0.2 per 1M tokens **Output tokens**: $0.4 per 1M tokens
Jamba 1.5 Large	The Jamba 1.5 Large model features 94 billion active parameters and 398 billion total parameters, with support for a 256K context window, allowing it to process up to 256,000 tokens at once. Built on the Jamba architecture, which combines Mamba and Transformer layers, it offers enhanced performance and efficiency for long-context tasks. The open model family democratizes access to LLMs that provide low latency, high quality, and robust long-context handling, enabling advanced GenAI applications like document summarization, text generation, and information extraction.	**Input tokens**: $2 per 1M tokens **Output tokens**: $8 per 1M tokens
AI21 Studio and other deployment options	AI21 Labs provides flexible deployment options for enterprises, including AI21 Studio for secure, production-ready integration, coud partnerships for adaptable AI solutions, AI21-managed private deployments in the customer's VPC with full data privacy, and self-managed options for complete control and compliance in on-premises or private cloud environments.	Custom pricing

GAI Insights

About GAI Insights

GAI Insights is the leading resource for news, research, and learning communities focused on Generative AI (GenAI), serving both companies and AI vendors. We offer year-long contracts with AI analysts who provide valuable market insights, best practices, a comprehensive Buyers Guide, and a database of successful use cases. Unlike other firms, we do not sell AI implementation projects. Our exclusive focus on AI and independence from vendors make us a trusted and impartial source for all things AI. Our insights have been featured in four articles in *Harvard Business Review* (HBR) and are highly regarded by our customer references.
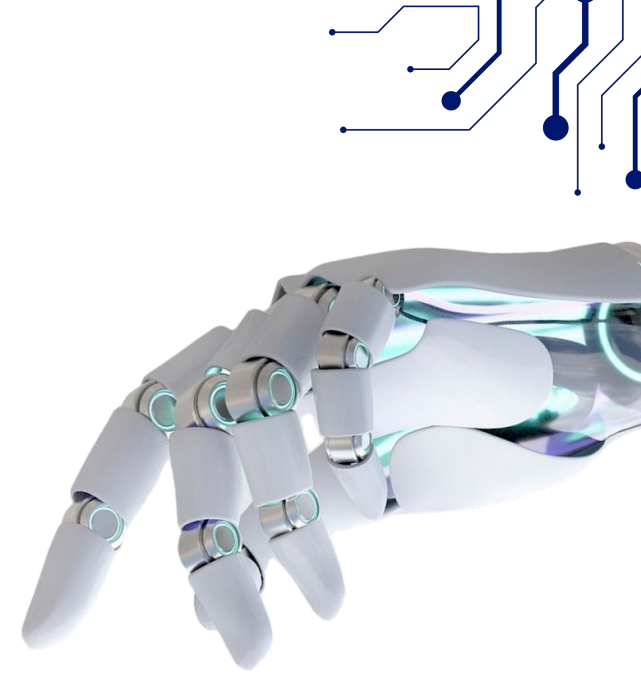
Stay Connected:

Explore more on our website: GAI Insights

Follow us on LinkedIn: GAI Insights LinkedIn

Read the blog articles: GAI Insights Substack

Follow us on Twitter: GAI Insights Twitter

Watch our YouTube: GAI Insights YouTube

GAI Insights

Like what you see?

Purchase the full guide for only $499 by scanning the QR code.

GAI Insights