

GAI Insights

# CORPORATE BUYERS' GUIDE TO LLMs

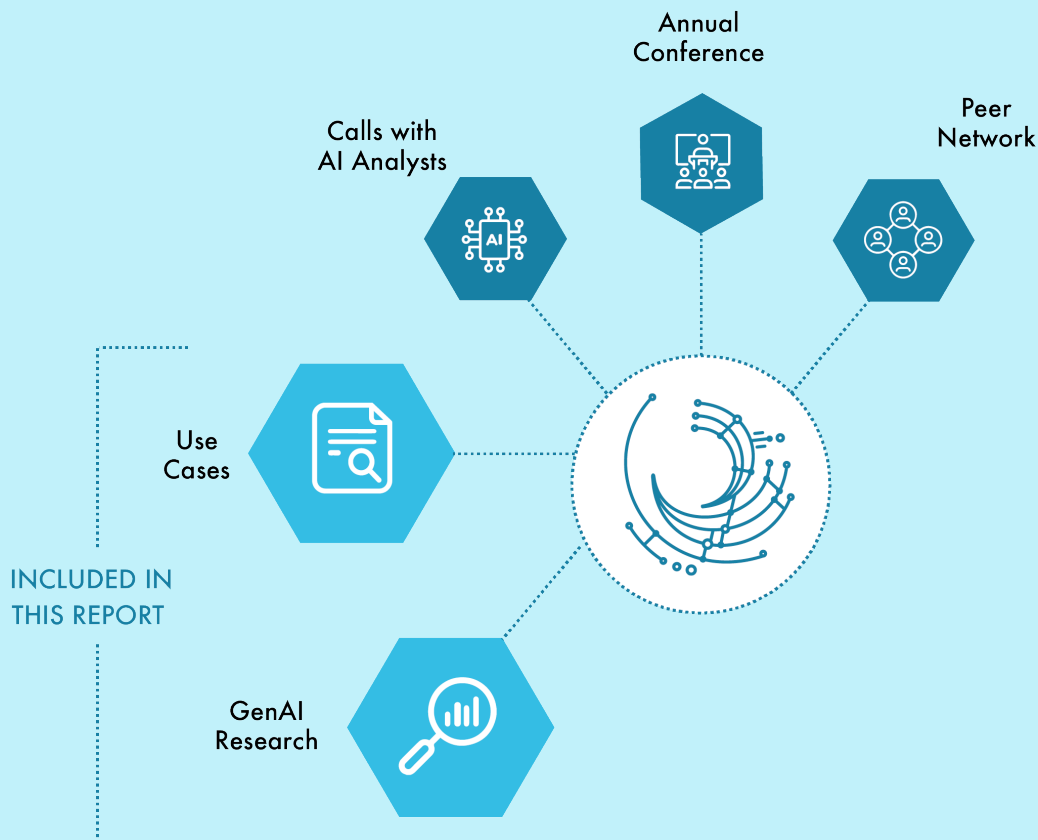
Defining and Answering Key Questions for  
Decision Makers in an Early Vendor Market



Q4 2023

# This Buyers' Guide is part of a yearly subscription for Premium Members

## GAI INSIGHTS PREMIUM MEMBERSHIP PROVIDES ACCESS TO:



## ABOUT GAI INSIGHTS

- A GenAI analyst firm dedicated to supporting AI leaders and their teams
- Comprises a team of leading AI Analysts
- Quantifies GenAI Applications through Use Cases
- Hosts an annual GenAI conference
- Fosters a community of over 3000 early GenAI adopters and enthusiasts
- Facilitates employee use of GenAI
- Maintains strong relationships with top academics from Harvard and MIT

**Harvard  
Business  
Review**

The WINS Framework, developed by Paul Baier (CEO at GAI Insights), Dr. John Sviokla (Co-Founder at GAI Insights), and Jimmy Hexter (Co-Founder at GAI Insights), has been published in the Harvard Business Review.

# SAVE THE DATE!

## Generative AI World 2024 Conference

Shift or Stagnate: Practical Strategies and Case Studies for Leaders

### October 7-8, 2024

The Intercontinental Hotel -- Boston, MA

### Learning - Networking - Sharing

Enter our ecosystem of conferences, research, meetups, learning labs, courses by joining us at our second annual conference. It is the only conference dedicated to enterprise case studies tailored to help the C-suite, AI leaders, board members and investors understand the promise and peril of this powerful, rapidly evolving, technology within the enterprise.

### HIGHLIGHTS

#### 20+ Enterprise GenAI Case Studies With ROI

- LLM Vendor selection
- Implementation best practices and horror stories
- Managing risks
- Best practices for organizational design, project prioritization, talent management

#### 45+ Speakers

- Corporate AI leaders
- AI Experts on law, etc.
- Academics from Harvard, MIT, Boston University and other experts.

#### Special Tracks For:

- Pharmaceutical and bio/pharmaceutical industry
- Financial services
- Healthcare
- Customer facing functions: sales, marketing and customer service

#### 15+ Vendor Exhibits With Focus On LLM Vendors

#### Fantastic Peer Networking

As a premium member of GAI Insights, you receive a 20% discount to the price of \$2,495.

### COMPANIES THAT ATTENDED IN 2023 TITLES

Autodesk  
Berkshire Partners  
Beth Israel  
Engage Smart  
Ensemble Health Partners  
Fidelity  
Gemline  
Howland Capital  
Houghton Mifflin Harcourt  
Liberty Mutual  
Mass General Brigham  
Mayo Clinic  
McGraw Hill Education  
PwC  
Takeda  
UBS  
WEX

Board Member  
CEO  
Chief Business Officer  
Chief Innovation Officer  
Chief Technical Officer  
Chief Digital Officer  
General Manager  
Managing Director  
Partner  
SVP AI



# Key Findings

*"We want our own ChatGPT with our data in our security framework, with **no hallucinations or errors.**"*

— CTO at a Fortune 500 Corporation

Generative AI (GenAI) is not a flash in the pan. It's poised to usher in a new way of interacting, working, and doing business akin to the advent of the Internet. Staying up-to-date in this nascent market is challenging, but understanding early market dynamics and developing a perspective on your company's unique opportunities lays the groundwork for a sustained competitive advantage.

1. **The business case for GenAI is very strong in many situations**, especially in industries and functions where significant value is generated through Words, Images, Numbers, and Sound (**WINS**) content, such as financial services, insurance, research and development, software engineering, customer service, and sales. GenAI leaders are hungry for a way to think about this rapidly transforming market, understand the specific case studies with quantified ROI, and implement best practices for business processes and technology management.
2. **Managing learning, adoption, and risk is vital.** Leading companies realize that GenAI is like swimming. You have to experience it to understand its value. Leaders invest in C-suite education, employee training and use policies, and small GenAI projects. Leaders also manage various risks around intellectual property (IP), bias, hallucinations, and security.
3. **The vendor market is alive with innovation, making early provider selection challenging.** \$27B has been invested in the LLM vendors covered in this Guide. A combination of major cloud providers, such as Google, Amazon, and Microsoft, and new market entrants, often with founders from big tech, such as Cohere and Anthropic, dominate the provider market. Significant competition exists between open source models, such as Llama from Meta, and proprietary models, such as GPT from OpenAI. Open-source is closing the performance gap. Careful vendor selection is key, and it's critical to avoid lock-in as market conditions, companies' priorities, and the competitive business landscape change rapidly.
4. **Based on our scoring criteria, we name ten LLM vendors as Emerging Market Leaders** (in alphabetical order): AI21 Labs, Aleph Alpha, Amazon, Anthropic, Cohere, CustomGPT, DataBricks, Google, Microsoft, and OpenAI. The products span no-code solutions for less than \$1K to ones over \$1M.

We are committed to helping AI leaders and their teams drive company value and competitive differentiation with GenAI. We value your feedback and welcome you to our GAI Insights member community.

— Paul Baier, Dr. John Sviokla, Echo, and the rest of the GAI team

**This Guide is a part of the  
yearly subscription service  
for GAI Insights Premium  
Members**

# Table of Contents

<b>Key Findings</b>	<b>3</b>
<b>About the Guide</b>	<b>6</b>
Who Should Read This Guide and Why	7
Research Methodology	9
About GAI Insights and Authors	12
<b>Part 1: MAKING THE BUSINESS CASE</b>	<b>15</b>
Using the WINS Framework to Assess GenAI Urgency and Value Potential	16
Case Studies with Quantified Benefits	20
GenAI is Here to Stay	24
<b>Part 2: LLM VENDOR SELECTION</b>	<b>25</b>
LLM Vendor Market Map	26
A Basic Overview of How LLMs Work	28
Five Evaluation Categories for LLM Vendors	29
Risk Management	31
Ten Steps to get Started with GenAI	33
Top High Tech Vendors Have vast Financial Resources to Invest in GenAI	36
Startups Raised \$27B	37
Ten Emerging Market Leaders	39
Vendor Directory	41
<b>Part 3: ADDITIONAL MATERIALS</b>	<b>73</b>
Key Concepts	74
The Historical Context: Beyond Ford and Taylor	78
How LLMs are Trained	83
Open-source versus Proprietary Models	86
Essential Policies and Procedures	90

# ABOUT THE GUIDE

The background features a dark teal gradient with vertical light streaks. In the foreground, a range of mountains is rendered as a glowing blue wireframe mesh. Numerous vertical lines of varying heights rise from the base of the mountains, each topped with a small, bright light. The background is also filled with a pattern of small, semi-transparent dots arranged in a grid-like structure, creating a sense of depth and digital space.

# Who Should Read This Guide and Why

This Guide is helpful for:

1. **Technology executives** who are **responsible** for using **GenAI** to drive the **technological growth and ROI**
2. **Corporate executives** who are **accountable** for understanding and **applying GenAI** in their firm for **tangible business results**
3. **C-suite and board members** interested in **comprehending** what **GenAI** is, how to **think** about it, and its **implications** for their **firm**
4. **Vendors** who want to help their clients **understand** the **GenAI landscape**
5. **Investors** seeking to **make sound decisions** on where to place their bets in an early and quickly **evolving vendor market**

## Why Read This Guide?

Key benefits include:

- A **business-strategy-oriented perspective** on GenAI
- **Actionable frameworks** to assess opportunities and accelerate adoption securely
- **Vendor evaluation criteria** reflecting key factors beyond technology capabilities
- Examples and **lessons** learned from **real-world case studies**
- **Guidance** on managing risks

More specifically, this Guide will provide the following value to leaders across functions and domains within an organization:



Leadership Role	Value
<b>Heads of AI, Analytics, and Digital Transformation</b>	<ul style="list-style-type: none"> <li>● Access tools to build business cases and evaluate vendors and licensing</li> <li>● Learn real-world case studies and funding data to inform your strategy</li> <li>● Discover best practices for governance, policies, and change management</li> </ul>
<b>Functional Leaders in Marketing, Sales, Service, HR, and Finance</b>	<ul style="list-style-type: none"> <li>● Identify high-potential GenAI applications in your domain</li> <li>● Quantify expected productivity gains from real-world examples</li> <li>● Learn how peers are piloting GenAI to transform operations</li> </ul>
<b>CEOs, CIOs, CTOs, CDOs, CISOs, and other C-suite leaders</b>	<ul style="list-style-type: none"> <li>● Get clarity on where GenAI can drive value and competitive advantage</li> <li>● Learn a strategic approach to evaluation, pilot projects, and scaling</li> <li>● Understand how to mitigate risks around data, ethics, and job impacts</li> </ul>
<b>Investors and Board Members</b>	<ul style="list-style-type: none"> <li>● Understand the technology and evolution of GenAI</li> <li>● Gain insight into market activity and assess investment opportunities</li> <li>● Appreciate risks and develop a prudent adoption roadmap for portfolio companies</li> </ul>

# Research Methodology

We present this Guide as a neutral resource. This Guide is not vendor-funded. It is unbiased and vendor inclusion is solely at our discretion.

To produce this Guide, we used the best of human intelligence and AI. Our human intelligence included interviews and survey research of practitioners, advisors, and technology suppliers in the GenAI space. Our AI tools included ChatGPT, Bard, Bing Chat, Perplexity, and Claude to assist in research, summarization, vendor analysis, and argument critiquing. We used Midjourney and DaLL-E 3 to create images.

Our **primary research** included:

- **In-depth discussions** with chief technology officers, chief digital officers, chief innovation officers, and other company leaders at 40 prominent companies
- **Briefings** with 15 leading vendors in the GenAI market
- **Consultations** with 12 respected consulting firms with expertise in GenAI
- **Conversations** with nine leading investors to gain insights into market trends and investments
- **Survey** among our 3,000-member GAI Insights community to capture a broad spectrum of perspectives
- **Two-day Generative AI World conference** with 300 attendees, 25 speakers, and showcasing 20 compelling case studies

Our **secondary research** involved using a set of GenAI tools for gathering, summarizing, and presenting information from research papers, blogs, and news outlets together with our analyst team. This helped our research and enabled our AI analysts to amass significant hands-on experience with using and, to some extent, building GenAI software.

## Companies Interviewed for the Guide

Below is a partial list of companies whose leaders and practitioners we interviewed:

### Companies

Atlantic Charter Insurance	Definitive Healthcare
Adventist Health	Electronic Arts
Clorox	Ensemble Healthcare Partners
CSX	Fidelity
Dualboot Partners	FleishmanHillard

Gainwell Technology	Northeast Color
Gemline	Omni Logistics
General Mills	Paramount
GlobalSource	Raytheon
Goodwin Procter	Right Networks
GridX	Schneider Electric
Houghton Mifflin Harcourt	Starr Insurance
Inizio Medical	Starnet Worldwide
Intel	Takeda
IP Australia	TE Connectivity
Jerry	Tomorrow.io
Liberty Mutual	Travelers Insurance
Lindt Chocolate	UPS
Mass General Brigham	Wayfair
Mayo Clinic	Wonderbotz
MKS Instruments	WEX
McGraw Hill Education	

### Industry Analysts

Eckerson Group	Latitude Media
Emerj	SanjMo
International Institute of Analytics	

### Private Equity Firms

American Securities	BV Investment Partners
Berkshire Partners	MidOcean Partners

### VC Firms

Battery Ventures	Glasswing Ventures
C10 Labs	Link Ventures
Flybridge Capital Partners	

## Consultants and Agencies

Bain and Company	FleishmanHillard
BigRio	KPMG
Boston Consulting Group	Ernst & Young
Coffee Labs	McKinsey
Digitas	PwC
Deloitte	Reliance Group

## Software Vendors

AI21 Labs	NetBase Quid
CapeStart	Pryon
Contextual.AI	Relevance AI
CRBNai	Roll
CustomGPT	Simplr
Lazarus AI	Snorkel AI
Microsoft	StackAware
Mistral.AI	Stochastic

REPORT EXCERPT

# About GAI Insights and Authors

GAI Insights is an analyst firm enabling AI leaders and their teams to achieve business results with GenAI. We specialize in enabling organizations to achieve remarkable ROI by strategically adopting GenAI technologies. We have established a reputation for being the first in the industry for several endeavors. For instance, we established the largest ChatGPT Meetup group and hosted the largest conference focused on enterprise GenAI case studies. Our insightful newsletters and daily news briefs are essential resources for leaders and investors across the globe.

## Our team



### PAUL BAIER

Paul is the CEO and principal analyst at GAI Insights. He is a seasoned software entrepreneur with two decades of experience. Paul has been an executive at multiple VC-backed startups, notably as VP of Product at First Fuel, an enterprise AI SaaS company. He holds an MBA from Harvard and a BA from Kenyon College.



### DR. JOHN SVIOKLA

Dr. Sviokla is the co-founder of GAI Insights. Throughout his career he has explored the practical implications of leading technologies. He is widely published and was a Partner at PwC, Vice Chairman of Diamond Technology Partners, and a Harvard Business School Professor, where he pioneered AI research and AI courses. Dr. Sviokla earned his Doctorate, Master's, and BA from Harvard University.



### TIM ANDREWS

Tim is an AI Analyst at GAI Insights. He is a luminary in innovative technology. Tim's past work includes a VP position at Booz Allen Hamilton, CTO at Viant Corp., and early software contributions to the Apple Macintosh. He holds an MSEE/CS from Worcester Polytechnic and a BA from Dartmouth.



## JULIO BARROS

Julio is an AI Analyst at GAI Insights. He is the founder of E-String Technologies, a cutting-edge AI consultancy. He holds a Master's from the University of Virginia and a BS from George Mason University.



## ADAM RAPPAPORT

Adam is an AI analyst at GAI Insights. He brings 30 years of experience across various industries, including finance, telecommunications, government, and emerging tech startups.



## AMANDA FETCH

Amanda is an AI Analyst at GAI Insights. She has two decades of experience in analytics, data science, AI/ML, and her work spans academia and corporate sectors. She is completing a PhD dissertation related to AI.



## VIVIAN PETERS

Vivian is an AI Research Analyst at GAI Insights. Vivian brings a solid decade of research experience. He holds a BBA from the University of New Brunswick.



## ECHO

Echo is the AI co-pilot of this Guide and a critical member of our team. Echo represents the cognitive intelligence of ChatGPT, Bard, Bing Chat, Midjourney, DALL-E 3, and other AI tools used for this Guide. It was a consistent and strong contributor.

We thank the following people for assisting us in producing and marketing this Guide: David DeLallo, Osama Elsayed, Pallavi Sud, Parika Kumar, Shameer Babu, Renan Amagan, Tania DiCostanzo, Chris Vander Rhodes, and Indigo Griffith as well as many other colleagues who reviewed earlier drafts of this Guide.

REPORT EXCERPT

PART I

# MAKING THE BUSINESS CASE





# Using the WINS Framework to Assess GenAI Urgency and Value Potential

There is an enormous amount of discussion in the media today around the potential long-term impact of GenAI. To best understand GenAI, we suggest looking at it through bifocal lenses. Looking through the top part of the lens, one sees the big, looming issues, such as accuracy, privacy, and bias. Firms such as Accenture and Goldman Sachs have written reports on the potential impact of GenAI on "knowledge workers" by industry and potential economy-wide job losses and societal risks. While this dialogue is important, we have heard repeatedly that it is not helpful to executives and board members. That is why we also recommend looking at GenAI through the bottom half of the bifocal lenses to find the immediate opportunities and threats.

## Why is GenAI Different? Because it is a Power tool for WINS Work

Our case studies based on our growing global community of 3,000 practitioners point to a new category of work, more precise and actionable than "knowledge work." We call it WINS work, where tasks, functions, and possibly your entire company or industry depend on manipulating and interpreting Words, Images, Numbers, and Sounds (WINS). Heart surgeons and chefs are knowledge workers, but not WINS workers. WINS workers are software programmers, accountants, marketing professionals, and many more.

## WINS = Words, Images, Numbers and Sounds

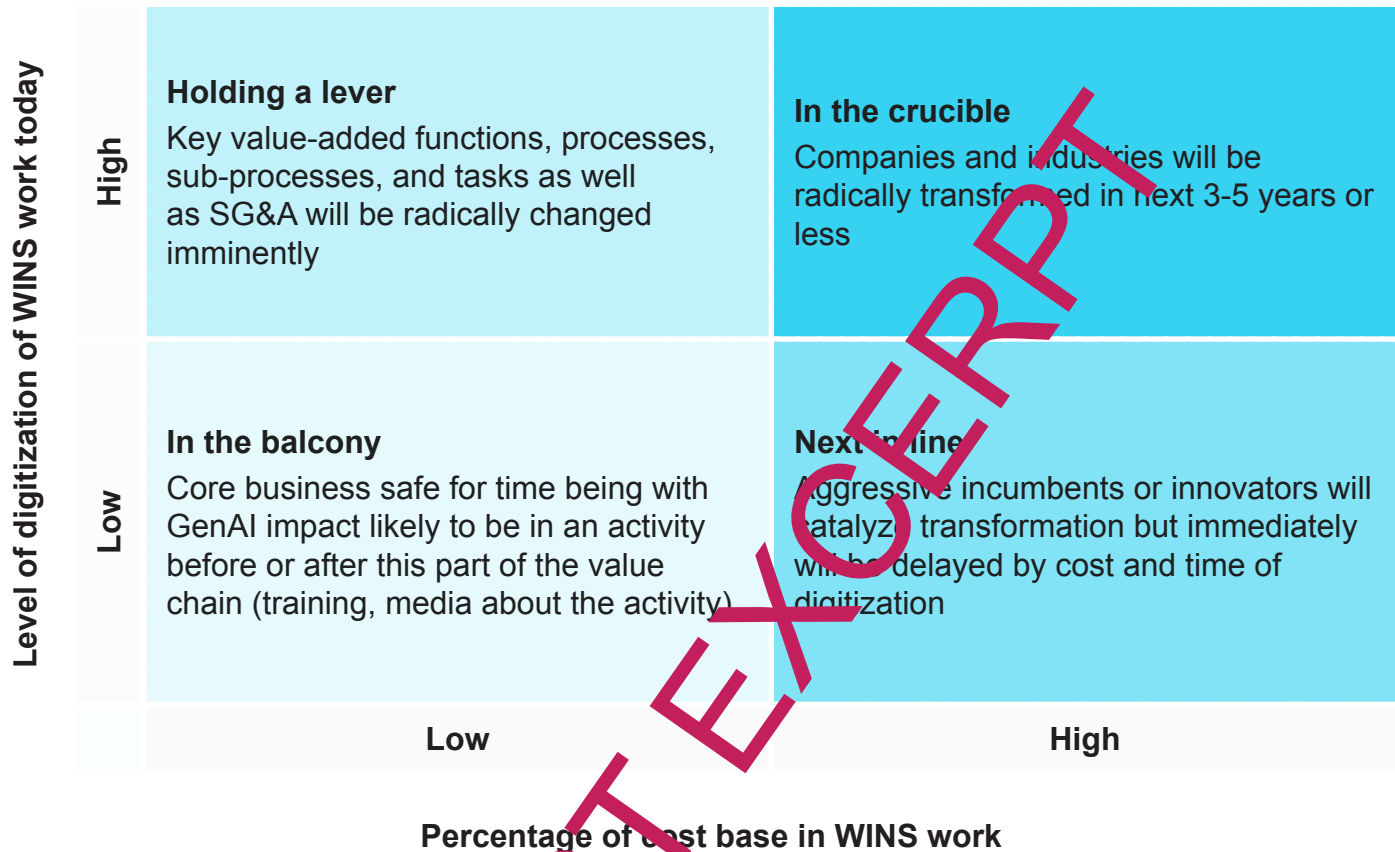
Think of GenAI as a power tool for WINS work. Would you hire a carpenter without a skilled saw or a roofer without a nail gun today? Every WINS task, sub-process, and end-to-end process within your enterprise, and in many cases the entire enterprise, should be evaluated for potential leverage with GenAI.

## How Urgent is it to Pay Attention to GenAI?

We believe the easiest way for companies to proceed is to ask themselves two simple questions:

- How much of our cost base is made up of WINS work?
- How digitized are the WINS inputs today?

## The WINS Framework



*An analysis of the level of digitization in various industries and their percentage of cost in the WINS work*

Industries with a high percentage of cost in WINS work and that are highly digitized are "**In the crucible**" and must understand and embrace GenAI immediately. The crucible includes the software industry, entertainment business, professional services, financial services, education, and others. Think of it like the demise of portrait painting. The camera made capturing of a likeness much easier and cheaper. Some portrait painters still exist, but they are few and far between. Coding and software development, script writing and film production, tax filing, and accounting will likely be under significant pressure. They may or may not become automated, but just as you'd never hire an accountant today who didn't use Excel, so too, an accounting firm in the future without their own GenAI capability will be less common, if not nearly non-existent.

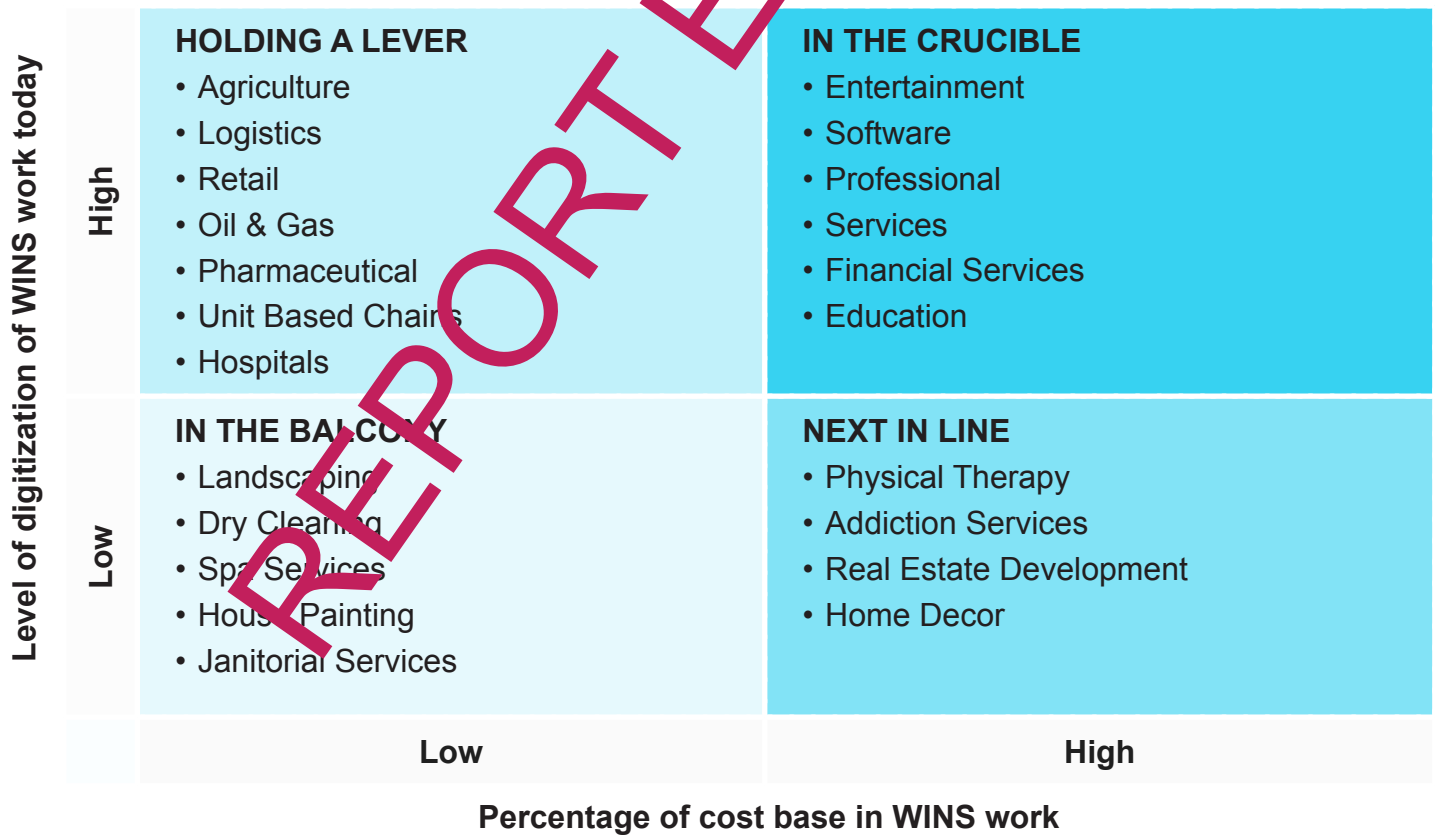
Companies that are "**Holding a lever**" can gain an advantage in cost, time, and quality even if their cost base is not heavily weighted toward WINS work and their product or deliverable falls within WINS and is digitized. For example, Moderna has just recently required that all employees be trained in GenAI tools. They believe it is a fundamental skill to drive WINS worker productivity, even though their product is a molecule or treatment intervention. In our GenAI learning community, we have found GenAI is excellent

for tasks such as supporting bid preparation when responding to a Request for Proposal. Speed of sale is a critical performance variable, and even those firms with few WINS workers could benefit by winning new work. Many selling, general, and administrative expenses functions, critical aspects of R&D, and even end-to-end product development and supply functions can leverage GenAI.

The "Next in line" category in our framework may allow us to take tasks that are not digitized today and digitize them to create opportunities. For example, many leading home décor companies are investing in what they are calling their digital front door, enabling customer engagement in the identification and purchase process. GenAI will allow new levels of customization to help customers take action to envision home furnishings in much more realistic and imaginative ways, leading to a better experience and greater customer uptake.

For companies that are "In the balcony," we see low digitization and limited WINS work as characteristic of the value creation process today. These are industries with high amounts of low-skilled labor, or when high skill is involved, and the nature of the skill is more in creating a physical product or service. The figure gives a sample of industries in each of the four quadrants.

**The WINS Framework with Industries**



*Examples of Industries in each of the four quadrants*



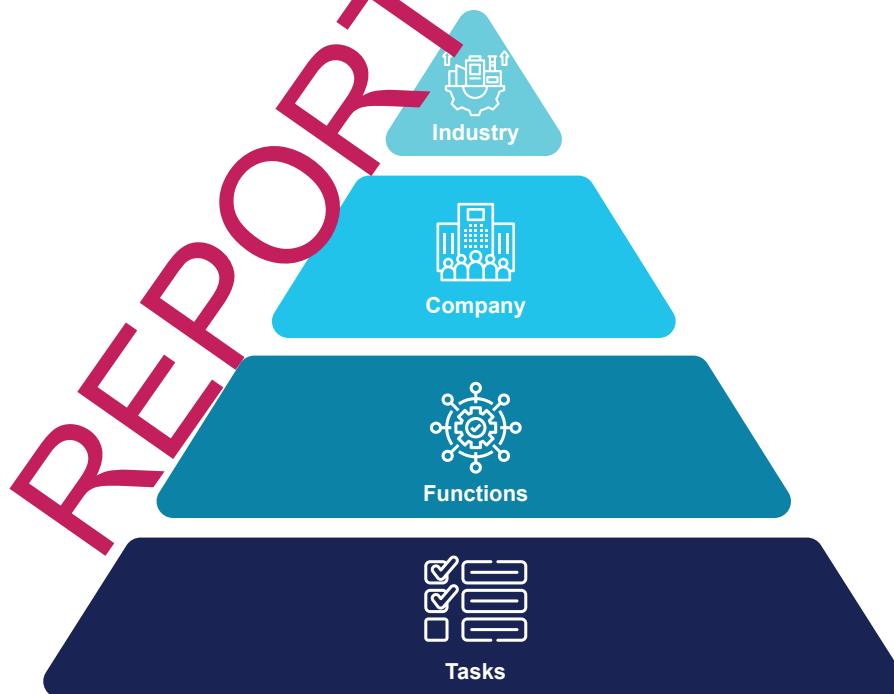
“What people have in their minds is this notion of tornados that come and wipe out an industry, ... it's going to be much **more like termites than tornados**.”

What you are going to hear is not the destruction of whole categories of work or categories of economic activity. Instead, you are going to see a gradual weakening of the business models

— Shikhar Ghosh, Professor at Harvard Business School, speaking at Generative AI World conference

Unlike many capital expenditures and technology spending that takes several years to see a return, GenAI, even at this early stage, can be accretive to earnings before interest, taxes, depreciation, and amortization (EBITDA) within the year it is adopted, in many cases, because the near-term productivity boost is so compelling. These initiatives may birth strategic investment opportunities to create defensible assets or competitive "moats."

## Look for Value from the Bottom Up



Organizations prioritize individual productivity at the Task level first, followed by Functional, Company, and Industry Levels

Our research found that most organizations start with individual productivity at the task level. From our community, we have sourced many individual case studies that exemplify this principle.

# Case Studies with Quantified Benefits

Below are three case studies that were taken from our learning community and our yearly Generative AI World conference in 2023:

**Case Study One:** GenAI customer care project is now helping save \$4M per year for the popular car insurance firm, Jerry. This GenAI project is already implemented and saving money. Jerry<sup>1</sup> is a company that aims to help consumers manage everything related to car ownership in one app, including comparing car insurance quotes, finding repair shops and quotes, and reaping savings through safe driving.

Since its inception in 2017, Jerry has experienced rapid growth, serving 5M customers. The company is fueled by \$110M in funding, received a \$450M valuation in 2021, and just earned a spot on Forbes' 2023 Best Startup Employers list. While the company is obviously thrilled about its success, its customer service model began experiencing stress that threatened to disenfranchise some customers.

At our inaugural Generative AI World Conference 2023 in Boston, Jerry's COO, John Spottiswood,<sup>2</sup> shared how using LLMs, the underpinnings of text-to-text GenAI, helped Jerry improve customer response times and save the company.



*The financial results of our customer service chatbot that is already in production has been simply spectacular with a 400% ROI and \$4M million in annualized savings.*



— John Spottiswood, COO of Jerry,  
speaking at Generative AI World 2023

## The Problem: Scaling Customer Service for Rapid Growth

With a focus on a fully digital and real-time experience, Jerry receives more than 200K messages a month from more than 100K unique users. The high volume meant the company struggled to respond to customer queries within 24 hours, let alone minutes or seconds.

<sup>1</sup> [Jerry | Company](#)

<sup>2</sup> [John Spottiswood | LinkedIn Profile](#)

PART II

# LLM VENDOR SELECTION

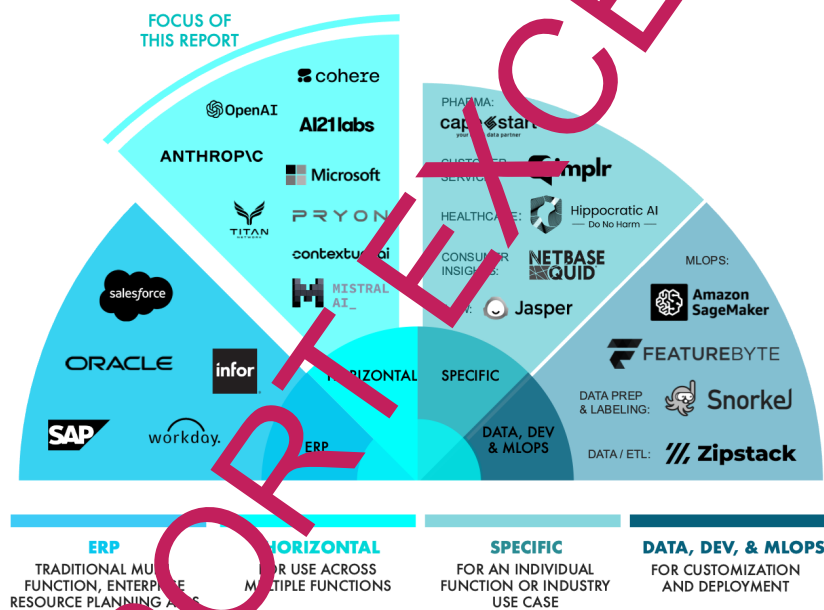


# LLM Vendor Market Map

This Guide focuses on generalized or horizontal LLM offerings that meet the use case needs for multiple company functions and processes.

The Guide categorizes vendors that:

- Sell proprietary LLM software related to GenAI deployments
- Use open-source or proprietary LLM models



*GAI Insights Market Map for LLMs though the scope of this Guide is on Horizontal LLMs*

The Guide does **not** cover the following:

- Implementation firms or consultants
- Enabling tools (e.g., Appen, LangChain) and enabling websites (e.g., Hugging Face)
- An exhaustive list of use cases, requirements, or a highly detailed LLM market sizing

This is a **vibrant and volatile vendor category**, with updates happening daily. Our list is current as of November 07, 2023. We used the following method to source our list:

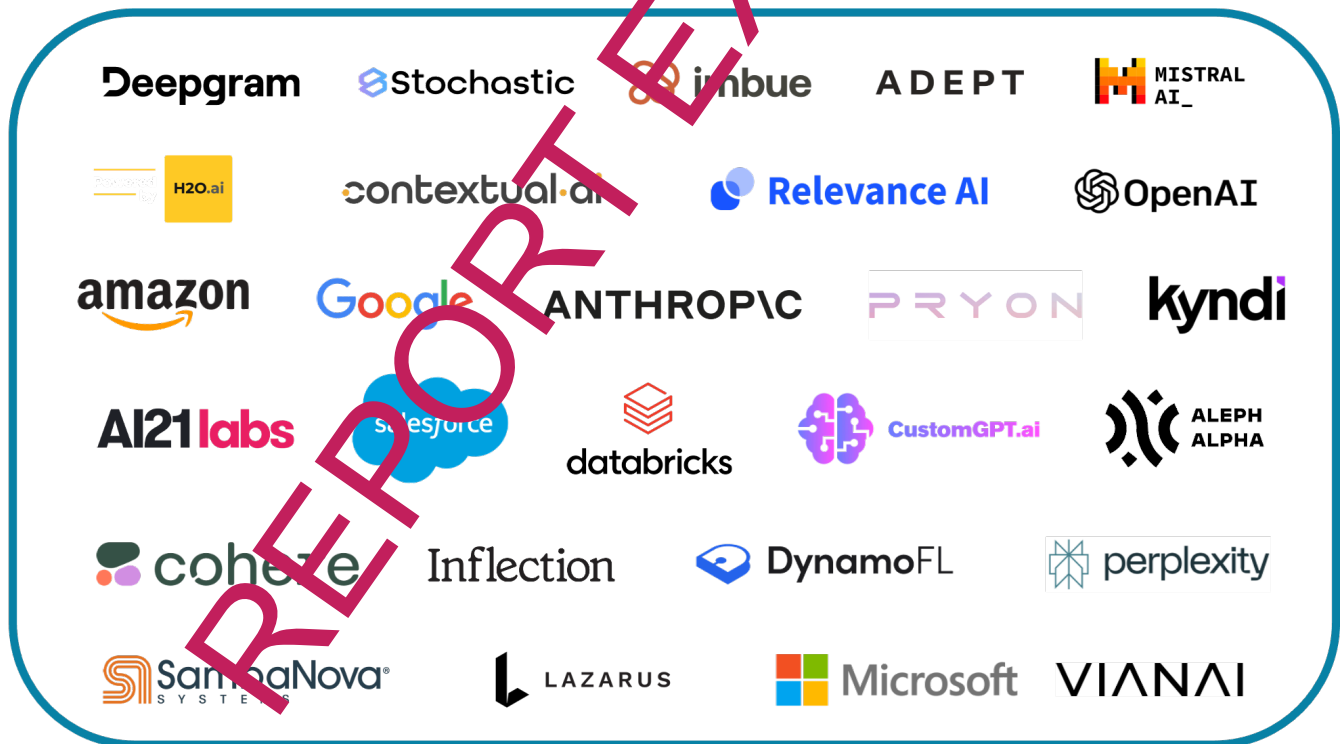
- Extensive use of GenAI and search tools
- Review of PitchBook and Crunchbase
- Analysis of leading AI publications, communities, and bloggers

- Consultation with our 3,000-user-strong Global AI community and investors
- Feedback and discussion at our Generative AI World Conference 2023

## LLM Vendors Evaluated

The precise definition of the LLM vendor category is changing, and it's debatable.

We started with an initial list of 80 vendors based on internet searches, articles, and feedback from the GAI Insights community. We added some and removed others as we learned more about each offering. Our guiding principle was providing help to CTOs or AI leaders facing requests for 5-15 LLM or chatbot projects in their organization from various departments, such as customer support, HR, and marketing. This Guide provides an unbiased evaluation. The vendors are listed based on our analysis. Despite our best efforts over the last four months, we undoubtedly missed a few and will address them in future updates.



*LLM vendors evaluated for this Guide*



# A Basic Overview of How LLMs Work

For this Guide, we will use the term LLMs. However, most LLMs are moving toward becoming large media models (LMMs) by integrating more functionality and data types. See a post by Dr. Sviokla<sup>3</sup> for additional discussion of this idea of LLMs becoming LMMs.

LLM<sup>4</sup> is the name of the underlying AI technology, and they are "generative" in the sense that they create or generate data, rather than simply predictions. ChatGPT, Claude, Bard, and BingChat are popular examples of LLMs with a browser interface, while applications can interact with the LLM via an API. With an LLM, a text prompt is entered, and a text response is "generated" and returned. With an LMM, the input can be text, audio, or an image, and the output can be a new image, new text such as a transcript or description, or new audio.



*I don't think machines will think like people – birds don't fly like airplanes. They fly differently, but they both fly. Machines think, people think – they're different.*

— Dr. John Sviokla, co-founder of GAI Insights, speaking at Generative AI World 2023

The term "language model" means a probabilistic model of language in modern use. Statistical models have been around, in theory, since the late 1940s, and in practice, they have been implemented in neural networks since the 2010s.<sup>5</sup> Note that by "language," we typically mean natural language, but it can be any coherent sequence of symbols such as code, music, or DNA sequences. Large LLMs are language models that have been scaled up to a previously infeasible scale with respect to both the model architecture and the data it is trained on.

<sup>3</sup> [Will All Large Language Models \(LLMs\) Become Large Media Models \(LMMs\)?](#)

<sup>4</sup> The tools evaluated in this guide, commonly called LLMs, have evolved from primarily handling text to encompassing images, numbers, and sounds. While they are transitioning to become Large Media Models (LMMs), we will use the more widely recognized term LLM for simplicity.

<sup>5</sup> [Recurrent Neural Networks and Long Short-Term Memory Networks: Tutorial and Survey](#)

# Five Evaluation Categories for LLM Vendors

In addition to the obvious fit in terms of company-specific use cases, IT infrastructure, and policies, we assert that five categories of vendor assessment are critical in this early market. These five categories formed the basis of our quantitative evaluation of all the vendors. Here is a discussion of each factor.

Category	Description	Evaluation	Relevance
<b>Product Innovation</b>	This refers to the introduction of new or significantly improved goods and services. It includes changes in design, user interaction, and technology used	Evaluate based on the uniqueness of the product or service, its improvement over existing solutions, and its potential impact on the market	Being first-to-market can confer advantages, but sustained success requires adaptation and fresh ideas. As Michael Porter (1996) <sup>6</sup> emphasized in his seminal work on business strategy, companies that continually innovate can maintain their competitive advantage and fend off rivals
<b>Quality of Investors</b>	This refers to the value the investor brings	Evaluate based on investors reputation and experience in AI, size of investor funds, historical investor performance, and alignment	In a rapidly developing market, key connections for introductions to sales prospects and new hires and additional capital (e.g., "dry powder") can materially make a difference for a startup
<b>Strategic Relationships</b>	These are mutually beneficial partnerships with other businesses, organizations, or investors. Complementary skills, customer relationships and other assets can speed growth	Evaluate based on the number of strategic relationships, their benefits, and their potential for future growth. Strategic relationships encompass alliances, partnerships, and collaborations	In GenAI, partnerships with academic institutions, tech companies, or industry-specific players can greatly amplify reach, impact, and growth. These relationships can offer shared resources, collaborative research, and access to new customer bases. Chesbrough (2006), <sup>7</sup> in his work on "Open Innovation," highlighted how external collaborations can enhance a firm's innovative prowess

<sup>6</sup> [What is Strategy?](#)

<sup>7</sup> [Beyond high tech: Early adopters of open innovation in other industries](#)

# Vendor Directory

## ADEPT

Adept sells a large-scale transformer model that can act as an AI assistant to browse, search, and use the web. The company raised \$415M from top-tier VCs and tech giants Microsoft and NVIDIA

URL	<a href="https://www.adept.ai/">https://www.adept.ai/</a>
Year Founded	2022
Headquarters	California, United States
Estimated Number of Employees	50 to 250
Total Funding	\$415M
Investors	General Catalyst, SVA, Frontiers Capital, Workday Ventures, NVIDIA, Microsoft
Product Lines	ADEPT, Model: Persimmon-8B, Multimodal Model: Fuyu-8B
Open-source / Proprietary	Open-source
Key Customers	Not disclosed



# AI21 labs

AI21 Labs offers task-specific LLMs for reading and writing. The firm raised \$283M from Google, NVIDIA, and others



URL	<a href="https://www.ai21.com">https://www.ai21.com</a>
Year Founded	2017
Headquarters	Tel Aviv, Israel
Estimated Number of Employees	50 to 250
Total Funding	\$283M
Investors	Walden Catalyst, Pitango, SCB10X, b2venture, Samsung Next, Google, NVIDIA
Product Lines	AI21 Studio, Models: Jurassic-2, Jurassic-X
Open-source / Proprietary	Proprietary
Key Customers	Carrefour, Clarivate, eBay, Guesty, Monday.com, ONE ZERO Bank, Ubisoft

A screenshot of the AI21 Studio website. The header includes the "AI21 studio" logo, navigation links for "Product", "Resources", and "Pricing", and buttons for "Log In" and "Start Building". The main content area features the headline "Differentiate your product with generative text AI" and a sub-headline "AI21 Studio provides API access to Jurassic-2 and Task-Specific language models. Our models power text generation and comprehension features in thousands of live applications." Below this are two buttons: "Start Building - It's Free" and "Let's Talk". On the right, there is a visual representation of a blog post editor with a "Blog Post Title" field and a code editor window showing Python code for using the AI21 API.

# Open-source versus Proprietary Models

Certain organizations may want to explore running their own model, possibly trained on their own data or hardware, primarily for privacy or competitive advantage reasons. This is certainly possible, though the associated costs of development and upkeep should be carefully considered. Along with the monetary costs, an organization will have to invest in finding which open-source or proprietary model best meets its needs.

The table below summarizes the examples of Open-source and Proprietary Models in the market: When choosing between an open-source LLM and a proprietary one for a business organization, there are several critical issues to consider:

Types of LLMs and associated Vendors' Models	
Open-source	Proprietary Models
<ul style="list-style-type: none"> <li>● BERT</li> <li>● BLOOM</li> <li>● Falcon</li> <li>● GPT-NeoX</li> <li>● Llava (an LMM)</li> <li>● Meta's Llama 2 and its variations (Platypus, Open-Orca, etc.)</li> <li>● Mistral 7B</li> <li>● OPT-175B</li> <li>● Zephyr 7B</li> </ul>	<ul style="list-style-type: none"> <li>● Anthropic Claude</li> <li>● Cohere</li> <li>● Google's Bison, Gecko, Gemini, and PaLM</li> <li>● OpenAI's GPT</li> </ul>

1. **Liability Risk:** With open-source models, liability risk may be lower as they are often provided "as is," and there is more control or choice over the data used for initial training. However, it's essential to understand the terms of the license agreement. Proprietary models might come with warranties or indemnities that can mitigate liability risk.
2. **Intellectual Property Concerns:** Open-source models often come with licenses that require any modifications to be made available under the same license. This could impact the proprietary nature of your business solutions. Proprietary models usually don't have this requirement. The major vendors, Microsoft, Google, and IBM, will indemnify their clients using their chat interface against IP claims when users follow guidelines and stay within the environment.
3. **Availability of Talent:** Open-source models often have a large community of users, making it easier to find skilled talent. On the other hand, proprietary models require more specialized knowledge. We also need to acknowledge vendors such as Amazon with the Amazon Bedrock Platform, IBM's

Watsonx.ai, and the NVIDIA AI platform, which offer developer tools such as MLOps Synthetic Data that assist in creating and deploying GenAI.

4. **Security:** Both types of models have security risks. Open-source models are transparent and can be reviewed by anyone, leading to early detection of vulnerabilities and potential exploitation. Proprietary models are not transparent, but they are supported by the vendor who is responsible for addressing security issues. Microsoft, IBM, Amazon, and others have built-in security and compliance.
5. **Ease of Implementation:** Open-source models can be more flexible and customizable, but this can also make them more complex to implement. Proprietary models might be easier to implement but could be less flexible.
6. **Cost:** Open-source models are usually free to use, but costs might be associated with implementation, maintenance, and support. On the other hand, proprietary models usually involve licensing fees.
7. **Support and Maintenance:** Open-source models often rely on community support, which can be inconsistent. Vendors of proprietary models typically offer consistent support and maintenance as part of the licensing agreement.
8. **Integration:** An open-source model may provide more flexibility and compatibility with your existing technology stack, as it can be modified to meet specific needs. However, a proprietary model may offer more seamless integration if the vendor has already developed solutions compatible with standard technology stacks.
9. **Scalability:** Open-source models can be adaptable to the growing needs of a business, but they may require more effort to ensure scalability. On the other hand, proprietary models often come with built-in scalability features, but at a potentially higher cost. As your organization expands, it's crucial to evaluate how well each model can handle increased data processing needs.
10. **Updates and Improvements:** With open-source solutions, updates might be frequent but inconsistent, while proprietary vendors typically provide regular updates. The most successful products are those that continue to innovate. Therefore, when evaluating an open-source or proprietary model, it's essential to gauge their commitment and capacity for improvement.

## Other Important Considerations

In addition to the above factors, we want to highlight other considerations your organization must take into account while narrowing down on an LLM vendor. This section helps you answer: **What should you consider when choosing an LLM vendor specific to your business application and feasibility?** Think of these pointers as benchmarks against which to streamline your vendor choice process.

1. **Model Performance:** Model size is essential, especially when considering the quality of answers. All else being equal, larger models will generate better answers, but will be more challenging to deploy. It may also take longer to process requests. Additionally, the amount, type, and quality of data used in pre-training and fine-tuning data are crucial to performance. Setting aside copyright