A Comparison of Generative AI Platform Offerings

Published 13 September 2023 - ID G00799592 - 4 min read

By Analyst(s): Justin Tung, Leinar Ramos, Arun Chandrasekaran, Haritha Khandabattu, Darin Stewart

> Generative AI technologies have proliferated throughout the technology industry. CIOs, application leaders and other IT leaders can use this downloadable presentation to better understand the market landscape for generative AI technologies, especially within the context of their own needs.

Beta Research

The following research is part of a new initiative Gartner is piloting to provide updates at a greater frequency. It is a work in progress that does not represent our final position. While we continue to monitor this topic, we invite you to provide constructive feedback. All relevant updates and feedback will be incorporated into the final research, which will undergo our standard review process.

Overview

Preface: This document currently includes a select number of vendors and is not an exhaustive list of all the competitors in this space. **More vendors and products were added on 13 September 2023.**

An increasing number of technology vendors are announcing and releasing a slew of generative AI (GenAI) products into the AI marketplace, providing organizations with a glut of options to choose from. This abundance of choice is confusing for CIOs and other IT leaders.

These presentation slides address the need for an aggregated market landscape view and comparison, and provide an early basis for contrasting the existing choices in the market.
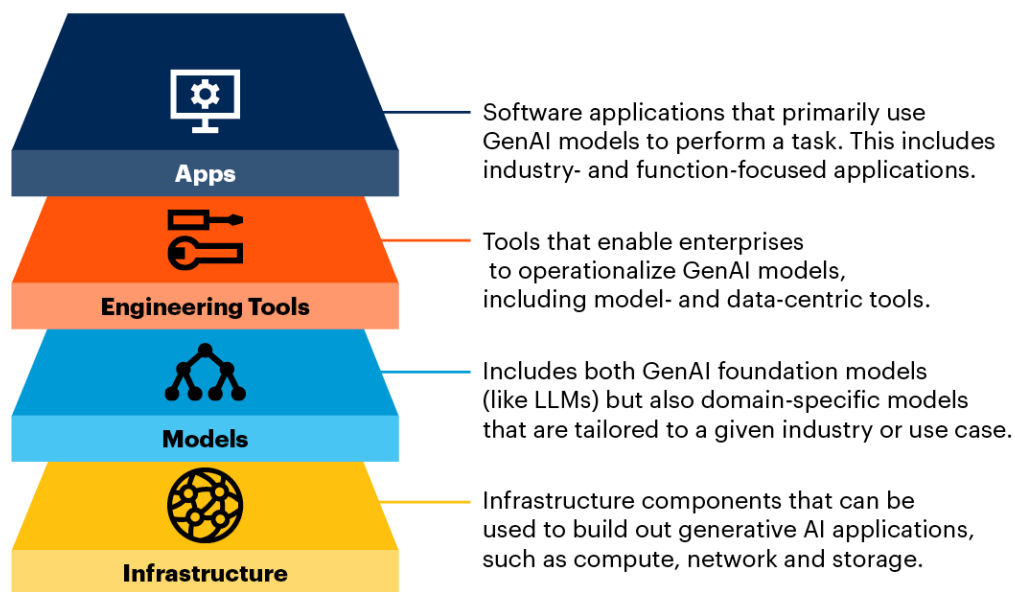
Download the Presentation Slides

GenAI is a multifaceted technology, with four different layers in its technology stack: GenAI applications, GenAI engineering tools, GenAI models and the underlying infrastructure to run these models (see Figure 1).

Many vendors offer solutions across these four technology layers, which increases the complexity in navigating the market. This downloadable deck maps the offerings of a limited number of GenAI platforms across these four layers, providing a clear way to map the market landscape and visually compare offerings.

Figure 1: Four Generative AI Technology Stack Layers

Four Generative AI Technology Stack Layers

Apps — Software applications that primarily use GenAI models to perform a task. This includes industry- and function-focused applications.

Engineering Tools — Tools that enable enterprises to operationalize GenAI models, including model- and data-centric tools.

Models — Includes both GenAI foundation models (like LLMs) but also domain-specific models that are tailored to a given industry or use case.

Infrastructure — Infrastructure components that can be used to build out generative AI applications, such as compute, network and storage.

Source: Gartner
799592_C

Gartner

Crucially, some organizations have a greater need for one of these layers over the others. For example, some enterprises may prefer to customize GenAI models hosted in their cloud infrastructure, while others may prefer to use an out-of-the-box GenAI application. Much of this is driven by business needs and resources — considering factors such as time to market, cost to implement, costs to deploy/maintain/monitor and the uniqueness of the unmet business need.

Utility

This is a tool for mapping the GenAI-related offerings by several major GenAI technology vendors. The market landscape comparison is conducted across several categories, including GenAI applications, engineering tools, models and infrastructure.

This comparison is **not** meant to be a ranking system, but simply a descriptive artifact through which one can easily view options for GenAI solutions. A vendor's lack of entries in a specific category does not necessarily mean that specific vendor is lagging behind; it may simply indicate that category is not a specialized area of focus for the vendor.

Note that some vendors have partnerships to offer models or capabilities that belong to another vendor. In addition, more than one vendor may share the same external partnership with another vendor. The usage of some offerings or models may also have restrictions. Always check vendor and offering/product usage policies. Some products might be early-access only, with general availability coming out later.

When designing your solution architecture, be aware that these GenAI offerings may need additional components that are not explicitly listed here. For example, components such as data warehouses or databases may be needed within your infrastructure layer.

Finally, this tool does not cover the pricing plans offered by individual vendors. However, with any vendor or offering, consider how costs are structured and whether you're charged by token submission, number of active users or a hybrid of the two.

Acronym Key and Glossary Terms

Data fabric	A design concept that serves as an integrated layer (fabric) of data and connecting processes. A data fabric uses continuous analytics over existing, discoverable and inferenced metadata assets to support the design, deployment and use of integrated and reusable data across all environments, including hybrid and multicloud platforms.
Foundation models	Large machine learning models/AI neural networks trained on a significant amount of raw (often unsupervised/unstructured) data.
Generative AI application	Any software application that primarily uses a generative AI function to perform a task or specific capability.
Generative AI engineering tool	An interface that makes it easier/more efficient to build generative-AI-powered artifacts.
Generative AI infrastructure	Infrastructure components that can be used to build out generative AI applications, tools, etc.
Generative AI model	(also see foundation models) Machine learning models that can be adapted to specific use cases for better performance in certain areas.
GPU	Graphics processing unit

Note 1: Vertical and Horizontal Generative AI Applications

Some of the products being offered by these vendors may be vertically or horizontally oriented.

Vertical applications focus primarily on specific industries and domains. Generative AI applications and offerings within this category are typically built on workflows and data most often seen within the space, such as pharmaceutical and life science research, legal and engineering design.

Horizontal applications are not necessarily tailored specifically to one industry or domain, but rather, they reach across different tasks and functional areas. These include products such as code generation tools, creative tools and HR task-oriented tools.

To determine which type of application might be best for your use case, consider whether your needs are highly specialized and specific to your industry. If so, a vertical application could be a good option. If your needs are more specific to the task or function you would like to perform, horizontal applications are a viable approach.

Recommended by the Authors

Some documents may not be available as part of your current Gartner subscription.

How to Pilot Generative AI

Innovation Guide for Generative AI Technologies

Quick Answer: How Do I Compare LLMs?