# Enterprises embrace model gardens as multi-model AI ecosystems accelerate



Enterprises are rapidly transitioning from reliance on a single large language model (LLM) provider to embracing what is known as the "model garden" approach, where multiple AI models from different vendors are utilised selectively to optimise performance across diverse tasks. This shift, accelerated over the past year, reflects the growing complexity and specialised nature of real-world AI applications, alongside the maturation of hosting layers such as Amazon Bedrock, Microsoft Azure AI, and Google Vertex AI. These platforms now provide unified APIs that enable seamless access to multiple model families, guardrails, observability, and retrieval capabilities, facilitating the deployment of multi-model assistants and agents that dynamically route requests to the best-performing model per task.

The fragmentation in model behaviour is a key driver of this change. Different models excel in varying aspects—some better at tone and empathy, others in domain-specific reasoning or handling long contextual inputs. For instance, Atlassian has developed an “AI Gateway” that orchestrates over 20 models from providers like OpenAI, Anthropic, and Google, dynamically routing queries based on accuracy, cost, latency, safety, and compliance requirements. Salesforce has similarly broadened its AI partnerships, incorporating OpenAI’s GPT-5 and Anthropic Claude into its recently launched Agentforce 360 platform, aiming to integrate generative AI tools across sectors including finance and healthcare, where regulatory controls are paramount.

Retail giant Walmart illustrates a textbook multi-provider deployment. In 2024, it introduced Wallaby, a retail-specific family of LLMs trained on decades of proprietary data, designed to be combined with other models for nuanced, contextualised responses. By October 2025, Walmart deepened its AI capabilities by partnering with OpenAI to enable customers to make purchases directly via ChatGPT’s Instant Checkout feature, blending strengths across models to optimise tone, latency, cost, and user experience in real time.

The need for compliance, data sovereignty, and localisation further reinforces this multi-model strategy. Enterprises like Vodafone split workloads geographically and by function—Azure OpenAI services handle customer assistant experiences, while Google Cloud manages network analytics and security operations. SAP’s Generative AI Hub integrates models from various providers including Amazon Bedrock and IBM watsonx Granite, granting customers greater choice and data sovereignty within one enterprise platform.

Cost pressures also promote continuous evaluation and optimisation. Companies such as Showpad and Rexera have centralised their model choices via Amazon Bedrock’s single API access to Anthropic Claude and Meta Llama models, achieving measurable cost efficiencies and improved latency. Financial services are adopting robust multi-cloud strategies with frameworks like the Fintech Open Source Foundation’s Common Controls for AI Services (CC4AI), which supports enterprise-grade governance across hybrid AI model deployments. Visa exemplifies this approach by leveraging a mix of models from OpenAI, Anthropic, IBM, Mistral, and Meta Llama, balancing open and closed models according to workload sensitivity and regulatory demands.

Technological advances in hosting layers simplify these complex multi-model ecosystems. Amazon Bedrock recently announced the general availability of multi-agent collaboration capabilities, enabling enterprises to build scalable AI workflows where multiple AI agents coordinate to perform sophisticated, multi-step tasks. Microsoft Azure AI Foundry integrates providers including OpenAI and Mistral, with enhanced security controls such as Prompt Shields to protect enterprises’ intellectual property and maintain compliance. Google Vertex AI offers a Model Garden across hundreds of models with guardrail options and agent tools, supporting experimentation and flexible deployment.

Security and governance remain paramount in this evolving landscape. Microsoft Defender for Cloud now provides AI Security Posture Management (AI-SPM) that spans AI workloads across Azure OpenAI, Amazon Bedrock, and Google Vertex AI, delivering vulnerability discovery, attack path analysis, and actionable recommendations to mitigate risks in multi-model, multi-cloud environments. Enterprises are advised to implement stringent guardrails, data residency policies, latency and cost monitoring, and continuous evaluation mechanisms. Deployments often start with policy-driven dynamic routing rather than hard-coded logic, enabling agile adaptation to emerging model releases and shifting workload demands without disruptive redeployments.

For organisations contemplating their AI strategy, the emerging decision framework entails mapping workloads by modality, complexity, compliance, and cost parameters, followed by selecting candidate providers for each task. Automated evaluation harnesses—leveraging ground truth scoring, human preference modelling, red-team testing, and telemetry—inform ongoing routing decisions. Hosting platforms should align with existing enterprise cloud infrastructure to facilitate observability and control, while governance frameworks like FINOS CC4AI enable consistent compliance across vendor and cloud boundaries.

Nonetheless, scenarios remain where a single-provider model approach is sensible—especially where legal, security, or operational constraints limit vendor diversity, or workloads are narrowly defined and stable. However, this choice requires regular reassessment as AI capabilities evolve and usage scales.

The model garden era represents a positive, pragmatic evolution in enterprise AI adoption, enabling organisations to harness the specialised strengths of diverse AI models to produce safer, more cost-effective, and higher-quality outcomes. It marks an important step towards building a UK and global AI environment characterised by responsible innovation, enhanced compliance, and optimised real-world value. As major players like Salesforce, Walmart, Visa, and DoorDash demonstrate, multi-model AI systems powered by advanced hosting platforms are increasingly becoming the new norm in AI-driven business transformation.

Source: [Noah Wire Services](https://www.noahwire.com)

## Bibliography

1. <https://medium.com/@graison/the-model-garden-era-choose-the-best-model-for-each-job-with-real-world-use-cases-b7587c2d4cc3?source=rss------machine_learning-5> - Please view link - unable to able to access data
2. <https://www.reuters.com/business/salesforce-deepens-ai-ties-with-openai-anthropic-power-agentforce-platform-2025-10-14/> - Salesforce has expanded its partnerships with OpenAI and Anthropic to integrate their advanced AI models—OpenAI’s GPT-5 and Anthropic’s Claude—into its newly launched Agentforce 360 platform. This initiative aims to enhance the platform as a hub for enterprise-grade generative AI tools across various business sectors, including regulated industries like finance and healthcare. The integrations will enable users to interact with customer data and analytics via tools such as ChatGPT, Slack, and Salesforce's own software. Additionally, a new feature called 'Agentforce Commerce' allows merchants to sell products using ChatGPT’s Instant Checkout while maintaining control over customer data and order fulfillment.
3. <https://www.reuters.com/business/salesforce-launches-agentforce-360-ai-platform-boost-software-products-2025-10-13/> - Salesforce has globally launched its AI platform 'Agentforce 360' across its suite of cloud-based tools to help clients automate routine tasks. The platform is designed to connect humans, AI agents, and data on a single, trusted system, enhancing productivity and efficiency. Agentforce 360 already has 12,000 customers, including major companies like Reddit, OpenTable, and Adecco. As competition grows in the AI-driven software industry, firms are increasingly integrating such technologies to help clients cut costs. Additionally, Salesforce revealed that its messaging tool Slack will incorporate conversational AI, enabling employees to retrieve data and perform tasks within chats, while maintaining enterprise-level security and controls.
4. <https://medium.com/@adtanasa/multi-cloud-multi-model-generative-ai-apps-with-azure-openai-vertex-ai-and-amazon-bedrock-73c2a90faca5> - This article discusses the integration of multiple cloud services—Azure OpenAI, Google Vertex AI, and Amazon Bedrock—to build multi-cloud, multi-model generative AI applications. It highlights the models available on each platform, such as GPT-4o and DALL·E on Azure, Gemini Pro/Flash and Imagen on Google Cloud, and Anthropic Claude and Meta Llama on Amazon Web Services. The piece also covers cross-cloud and intra-cloud connectivity options, including Azure's APIM with Azure OpenAI, Google Cloud's Workload Identity Federation with Vertex AI, and Amazon Bedrock's serverless deployment capabilities. The article provides insights into leveraging these platforms for scalable and cost-effective AI solutions.
5. <https://www.cloudoptimo.com/blog/amazon-bedrock-vs-azure-openai-vs-google-vertex-ai-an-in-depth-analysis/> - This in-depth analysis compares Amazon Bedrock, Azure OpenAI, and Google Vertex AI, focusing on their model access types, execution modes, and suitability for various use cases. Amazon Bedrock offers a multi-vendor approach with a fully serverless, API-based execution mode, ideal for testing multiple vendors quickly. Azure OpenAI provides consistent GPT access at scale with Azure-hosted, provisioned endpoints, suitable for GPT-focused enterprise workloads. Google Vertex AI offers flexibility and experimentation with a managed or user-deployed execution mode, supporting both proprietary and open-source models within the same ML pipeline. The article helps organizations choose the right platform based on their specific needs.
6. <https://aws.amazon.com/blogs/machine-learning/amazon-bedrock-announces-general-availability-of-multi-agent-collaboration/> - Amazon Bedrock has announced the general availability of multi-agent collaboration, allowing developers to build, deploy, and manage networks of AI agents that work together to execute complex, multi-step workflows efficiently. This capability has been utilized across various industries, including financial services, healthcare, supply chain and logistics, manufacturing, and customer support. The release introduces enhancements based on customer feedback, improving scalability, observability, and flexibility, making AI-driven workflows easier to manage and optimize.
7. <https://techcommunity.microsoft.com/blog/microsoft-security-blog/enhance-ai-security-and-governance-across-multi-model-and-multi-cloud-environmen/4395593> - Microsoft Defender for Cloud has expanded its AI Security Posture Management (AI-SPM) to support AI workloads across Azure OpenAI Service, Azure Machine Learning, and Amazon Bedrock. This enhancement includes support for Google Vertex AI models and improved compatibility with Azure AI Foundry model catalog and custom models. The expansion offers discovery of AI inventory and vulnerabilities, attack path analysis, and recommended actions to address risks in Google Vertex AI workloads, thereby enhancing AI security and governance across multi-model and multi-cloud environments.