



[知识库](#) > [Using Deskpro](#) > [AI](#) > [Deskpro AI Supported Models & Providers](#)

Deskpro AI Supported Models & Providers

Lara Proud - 2026-01-30 - [Comments \(0\)](#) - [AI](#)

Deskpro AI supports a wide range of **Large Language Models (LLMs)** and **Embedding Models** across multiple providers, so you can choose the right option for your organization.

We've tested these models extensively to ensure compatibility for both **prompting** (LLM-powered tasks like summarization and reply suggestions) and **embeddings** (semantic search and data retrieval).

This article lists all currently supported public and private models, along with guidance for supported AI models for Deskpro Cloud and Private.

Understanding Prompting and Embedding Models

Prompting Model

A prompting model is like a translator for your request — it turns what you ask into clear, step-by-step instructions that the LLM can follow.

Example: You say, "Write a friendly email to reschedule my dentist appointment." The prompting model reformats that request with tone and key points, and the LLM writes the email.

Embedding Model

Embeddings act like "meaning fingerprints" for text. An embedding model converts your words and documents into numeric vectors, allowing the LLM to find relevant information even when different words are used.

Example: You ask, "How do I get paid time off?" The embedding model retrieves the HR document that mentions "annual leave" and passes it to the LLM, so you get the correct policy in plain language.

Deskpro Cloud

On Deskpro Cloud, you can choose from a range of public AI providers and configure your account directly in Deskpro, or use the Deskpro Managed AI Connection to get started instantly.

Provider	Vendor	Model	Model Name	Type
OpenAI	OpenAI	Chat GPT 5.1	gpt-5.1	Prompting
OpenAI	OpenAI	Chat GPT 5	gpt-5	Prompting
OpenAI	OpenAI	Chat GPT 5	gpt-5-mini	Prompting
OpenAI	OpenAI	Chat GPT 5	gpt-5-nano	Prompting
OpenAI	OpenAI	Chat GPT 4.1	gpt-4.1	Prompting
OpenAI	OpenAI	Chat GPT 4.1	gpt-4.1-mini	Prompting
OpenAI	OpenAI	Chat GPT 4.1	gpt-4.1-nano	Prompting
OpenAI	OpenAI	Chat GPT4o	gpt-4o-nano	Prompting
OpenAI	OpenAI	Chat GPT4o	gpt-4o-mini	Prompting
OpenAI	OpenAI	Text Embedding	text-embedding-3-small	Embedding
OpenAI	OpenAI	Text Embedding	text-embedding-3-large	Embedding
Azure	OpenAI	Chat GPT 5	gpt-5	Prompting
Azure	OpenAI	Chat GPT 5	gpt-5-mini	Prompting
Azure	OpenAI	Chat GPT 5	gpt-5-nano	Prompting

OpenAI	OpenAI	Chat GPT 4.1	gpt-4.1	Prompting
OpenAI	OpenAI	Chat GPT 4.1	gpt-4.1-mini	Prompting
OpenAI	OpenAI	Chat GPT 4.1	gpt-4.1-nano	Prompting
OpenAI	OpenAI	Chat GPT4o	gpt-4o-nano	Prompting
Azure	OpenAI	Chat GPT4o	gpt-4o-mini	Prompting
Azure	OpenAI	Text Embedding	text-embedding-3-small	Embedding
Azure	OpenAI	Text Embedding	text-embedding-3-large	Embedding
Deskpro Managed Service	OpenAI	Chat GPT 5.1	gpt-5.1	Prompting
Deskpro Managed Service	OpenAI	Chat GPT 5	gpt-5	Prompting
Deskpro Managed Service	OpenAI	Chat GPT 5	gpt-5-mini	Prompting
Deskpro Managed Service	OpenAI	Chat GPT 5	gpt-5-nano	Prompting
Deskpro Managed Service	OpenAI	Chat GPT 4.1	gpt-4.1	Prompting
Deskpro Managed Service	OpenAI	Chat GPT 4.1	gpt-4.1-mini	Prompting
Deskpro Managed Service	OpenAI	Chat GPT 4.1	gpt-4.1-nano	Prompting
Deskpro Managed Service	OpenAI	Chat GPT4o	gpt-4o-mini	Prompting
Deskpro Managed Service	OpenAI	Chat GPT4o	gpt-4o-nano	Prompting
Deskpro Managed Service	OpenAI	Text Embedding	text-embedding-3-small	Embedding
Deskpro Managed Service	OpenAI	Text Embedding	text-embedding-3-large	Embedding
Anthropic	Anthropic	Claude 3.5 Haiku	claude-3.5-haiku	Prompting
Voyage	Voyage	Voyage 3.5	voyage 3.5	Embedding
Mistral	Mistral	Mistral Medium 3.1	magistral-medium-2507	Prompting
Mistral	Mistral	Mistral Medium 3.1	mistral-embed	Embedding
Vertex	Google	Gemini	gemini-2.0-flash-lite-001	Prompting
Vertex	Google	Gemini	gemini-embedding-001	Embedding

Deskpro Private

With Deskpro Private, you can either:

- Subscribe to a public AI provider (any listed above), or;
- Configure Deskpro to use your own self-hosted models through providers list below.

Provider	Vendor	Model Family		Model Name	Type
Ollama	OpenAI	GPT OSS	gpt-oss		Prompting
Ollama	Meta	llama 4	llama4		Prompting
Ollama	Meta	llama 3.3	llama3.3		Prompting
Ollama	Meta	llama 3.2	llama3.2		Prompting
Ollama	Meta	llama 3.1	llama3.1		Prompting
Ollama	Meta	llama 3	llama3		Prompting
Ollama	Meta	llama 2	llama2		Prompting
Ollama	Alibaba	qwen3	qwen3		Prompting
Ollama	Alibaba	qwen2.5	qwen2.5		Prompting
Ollama	Alibaba	qwen3	qwen3-embed		Embedding
Ollama	Mistral	mistral-large	mistral-large-2411		Prompting
Ollama	Mistral	Mistral Nemo 12b	open-mistral-nemo-2407		Prompting
Ollama	Mistral	Mistral 24.11	mistral		Prompting
Ollama	Mistral	Mixtral 8x7B	mixtral		Prompting
Ollama	Microsoft	Phi-4	phi4		Prompting
Ollama	Microsoft	Phi-3.5	phi3.5		Prompting
Ollama	Google	Gemma 2	gemma2		Prompting
Ollama	Google	Gemma 3n	gemma3n		Prompting
Ollama	Google	Gemma 3n	gemma3		Prompting
Ollama	Google	Gemma	embeddinggemma		Embedding
Ollama	Deekseek	Deepseek R1	deepseek-r1		Prompting

Ollama	Deekseek	Deepseek V3.1	deepseek-v3.1	Prompting
Ollama	Nomic AI	gpt4all	nomic-embed-text	Embedding
Ollama	Mixedbread	Mixedbread	mxbai-embed-large	Embedding
Ollama	BAAI	BGE-M3	bge-large	Embedding
Ollama	Microsoft	MiniLM	all-minilm	Embedding
AWS Bedrock	OpenAI	GPT OSS	openai.gpt-oss-20b-1:0	Prompting
AWS Bedrock	OpenAI	GPT OSS	openai.gpt-oss-120b-1:0	Prompting
AWS Bedrock	Anthropic	Claude Haiku 3.5	anthropic.claude-3-5-haiku-20241022-v1:0	Prompting
AWS Bedrock	Anthropic	Claude Haiku 4.5	anthropic.claude-haiku-4-5-20251001-v1:0	Prompting
AWS Bedrock	Anthropic	Claude Sonnet 4	anthropic.claude-sonnet-4-20250514-v1:0	Prompting
AWS Bedrock	Anthropic	Claude Sonnet 4.5	anthropic.claude-sonnet-4-5-20250929-v1:0	Prompting
AWS Bedrock	Meta	llama 4	meta.llama4-maverick-17b-instruct-v1:0	Prompting
AWS Bedrock	Meta	llama 4	meta.llama4-scout-17b-instruct-v1:0	Prompting
AWS Bedrock	Meta	llama 3.3	meta.llama3-3-70b-instruct-v1:0	Prompting
AWS Bedrock	Meta	llama 3.2	meta.llama3-2-90b-instruct-v1:0	Prompting
AWS Bedrock	Meta	llama 3.1	meta.llama3-2-11b-instruct-v1:0	Prompting
AWS Bedrock	Meta	llama 3	meta.llama3-1-70b-instruct-v1:0	Prompting
AWS Bedrock	Meta	llama 2	meta.llama3-1-8b-instruct-v1:0	Prompting
AWS Bedrock	Deepseek	Deepseek R1	deepseek.r1-v1:0	Prompting
AWS Bedrock	Amazon	Titan Embed	amazon.titan-embed-text-v1	Embedding
AWS Bedrock	Amazon	Titan Embed	amazon.titan-embed-text-v2:0	Embedding
AWS Bedrock	Amazon	Titan Text	amazon.titan-text-lite-v1	Embedding
AWS Bedrock	Amazon	Titan Text	amazon.titan-text-express-v1	Embedding

相关内容

- [Deskpro AI supports Anthropic Claude, Google Gemini, and Mistral AI](#)
- [Setting up Deskpro AI](#)