

LLM Comparison Tracker

Updated April 10, 2025

Chasing Next

BRAND	OpenAI	Google	Anthropic	xAI	DeepSeek	Alibaba	Meta	Perplexity AI*
PLATFORM	ChatGPT	Gemini	Claude	Grok	DeepSeek	Qwen	Llama	Perplexity*
URL	chatgpt.com	gemini.google.com/	claude.ai/	grok.com/	chat.deepseek.com/	chat.qwen.ai/	meta.ai/	perplexity.ai/
PREDICT	✔	✔	✔	✔	✔	✔	✔	✔
WEB SEARCH	✔	✔	✔	✔	✔	✔	✔	✔
DEEP RESEARCH	✔	✔	✖	✔	✖	✖	✖	✔
REASONING	✔ o1, o1-mini, o3-mini models	✔ 2.5 model	✔ 3.7 Sonnet model	✔ Grok3 model	✔ R1 model	✔ QwQ model	✔ Llama 4 Reasoning (Coming Soon)	✔ Reasoning Mode (DeepSeek R1, OpenAI)
VOICE CHAT	✔ Advanced Voice Mode	✔ Gemini Live via Gemini mobile app	✖	✖	✖	✔ Qwen2-Audio	✔ Llama 3.2	✖
IMAGE GENERATION	✔ 4x	✔ Imagen 3	✖	✖	✖	✔ Qwen 2.5	✔ Llama 3	✖
VIDEO GENERATION	✔ Sora	✔ Veo 2	✖	✖	✖	✔ Wan 2.1	✔ Movie Gen	✖
AUDIO GENERATION	✔ Text-to-Speech via API	✔ Chirp 3 via API	✖	✖	✖	✖	✖	✖
MUSIC GENERATION	✖	✔ Lyria (in preview via Vertex AI)	✖	✖	✖	✖	✖	✖
CAMERA VIEW	✔ Advanced Voice with Vision via mobile app	✔ Gemini Live via Gemini mobile app	✖	✖	✖	✖	✖	✖
MEMORY CAPABILITIES	✔ Persistent + conversational memory	✔ Persistent + conversational memory	✔ Conversational memory	✔ Conversational memory	✔ Conversational memory	✔ Conversational memory	✔ Conversational memory	✔ Conversational memory
CONTEXT WINDOW	✔ 4o, 4.5 = 128K tokens	✔ 2.5 Pro = 1M tokens	✔ 3.7 Sonnet = 200K tokens	✔ Grok3 = 1M tokens	✔ R1 = 128K tokens	✔ Qwen 2.5 Turbo = 1M, most others = 128K	✔ 4 Scout = 10M, 4 Maverick = 1M	✔ Default 4K tokens
OPEN SOURCE/WEIGHT	✖	✖	✖	✖	✔ R1, V3 = open weight	✔ 2.5, Qwen2-Audio = open source	✔ 4 models = open weight	✖
KNOWN KNOWLEDGE CUTOFF DATE	✔ o1, 4.5 = Oct 2023	✔ 2.5 Pro = Jan 2025	✔ 3.7 Sonnet = Oct 2024	✖ Grok3 unknown	✖ R1 unknown	✔ 2.5-Max = Dec 2024	✔ 4 models = Aug 2024	✖ Models vary
BEST FOR	Every Day, Deep Research, Image Generation	Reasoning, Video Gen	Writing, Creative Tasks, Coding	Every Day	Cost Effectiveness, Open Use	Open Use, Advanced Tasks	Open Use, Chatbots, Context Retrieval	Web Search

*Perplexity isn't an LLM, it uses models from ChatGPT, Claude, & DeepSeek behind the scenes. It's included because its search & research features are very useful, especially for real-time answers.

**Some models & features listed may only be available on paid plans or through APIs.

***There may be errors in the above chart. This landscape is changing fast—models improve, release new versions, or shift features often.