

AI engines operating today, with hyperlinks to their official or authoritative pages, highlighting their strengths and noting those with multimodal (text + image) capabilities:

AI Engine & Link	Developer	Strengths & Highlights	Multimodal (Text + Image) Capability
<u>GPT-4o</u>	OpenAI	Leading large language model with exceptional contextual understanding, multilingual support, and wide-ranging applications including code generation and creative writing.	Yes (via GPT-4o Vision)
<u>Claude 3.7 Sonnet</u>	Anthropic	Strong emphasis on ethical AI, safety, and nuanced reasoning. Excels in complex problem-solving and responsible AI deployment.	Limited multimodal; evolving
<u>Gemini 2.5 Pro</u>	Google DeepMind	Advanced contextual understanding, strong reasoning, and multimodal integration combining text, images, and other modalities.	Yes
<u>LLaMA 3.1</u>	Meta AI	Open-source with large parameter counts, adaptable across domains, strong community support, and efficient inference.	Emerging multimodal extensions
<u>Megatron-Turing NLG</u>	NVIDIA	High-speed, scalable text generation optimized for large-scale deployments, including narrative generation and interactive gaming.	Primarily text
<u>DeepSeek-V3</u>	DeepSeek	Very large model (671B parameters), open source, strong in multimodal search and retrieval, combining text and image processing.	Yes
<u>Mistral Large 2</u>	Mistral AI	Open-source, high-performance LLM with 123B parameters, optimized for efficient inference and diverse NLP tasks.	Primarily text
<u>PaLM 2</u>	Google	Strong in multilingual understanding, semantic search, and multimodal tasks; integrates well with Google ecosystem.	Yes
<u>Xlnet Pro</u>	Microsoft Research	Permutation-based learning model with superior generalization, excelling in translation and complex text analysis.	Primarily text
<u>Grok-3</u>	xAI (Elon Musk)	Focused on conversational AI with strong real-time interaction capabilities; evolving multimodal support.	Emerging multimodal
<u>Stable LM 2</u>	Stability AI	Open-source LLM with focus on democratizing AI access, strong community-driven development, and multimodal capabilities.	Yes
<u>Claude 3.5 Sonnet</u>	Anthropic	Prior version of Claude with strong ethical alignment and conversational abilities; widely used in enterprise.	Limited multimodal

<u>Nova</u>	Amazon	Cloud-optimized LLM for enterprise applications, scalable and secure with integration into AWS services.	Primarily text
<u>Qwen 3</u>	Alibaba	Large open API model with 235B parameters, strong in multilingual and cross-domain tasks.	Emerging multimodal

Notes on Multimodal AI Engines

- **GPT-4o Vision** and **Google DeepMind’s Gemini** series are currently the most advanced in integrating text and image processing, enabling richer human-computer interaction.
- **DeepSeek** specializes in combining large-scale text and image retrieval, making it powerful for search and content understanding.
- Many open-source models (e.g., LLaMA, Stable LM) are rapidly evolving toward multimodal capabilities, with community-driven extensions.

This table reflects the state of AI engines as of May 2025, synthesizing information from leading AI model leaderboards and industry reports^{[1][2][3][4][5][6][7][8]}.

1. <https://zapier.com/blog/best-llm/>
2. <https://www.techtarget.com/whatis/feature/12-of-the-best-large-language-models>
3. <https://www.shakudo.io/blog/top-9-large-language-models>
4. <https://artificialanalysis.ai/leaderboards/models>
5. <https://www.byteplus.com/en/topic/416048>
6. <https://explodingtopics.com/blog/list-of-llms>
7. <https://www.byteplus.com/en/topic/416052>
8. <https://www.empler.ai/blog/the-ultimate-guide-to-the-latest-llms-a-detailed-comparison-for-2025>