Comprehensive List of Large Language Models (LLMs)

1. Text-Only LLMs

Model Name	Primary Purposes	Release Year	Developer	Key Details
GPT-o3	Advanced reasoning, long-context tasks, code generation	2025	OpenAI	Closed-source; excels in math and multi-step planning; context: 128K tokens
GPT-4.1	General text generation, creative writing, Q&A	2025	OpenAI	Closed-source; optimized for efficiency over GPT-4o; context: 200K tokens
GPT-4.5	Multimodal precursor but text-focused; ethical alignment	2025	OpenAI	Closed-source; improved safety and bias reduction
Llama 4 Scout	Open-source text generation, fine-tuning for custom apps	2025	Meta AI	Open-source; 70B parameters; strong in multilingual tasks
Gemini 2.5 Pro	Reasoning, summarization, code; integrated with Google ecosystem	2025	Google DeepMind	Closed-source; context: 1M+ tokens; excels in long-form analysis
Claude 3.7	Ethical text generation, document analysis	2025	Anthropic	Closed-source; constitutional AI for safety
MiniMax-Text-01	Cost-effective text gen for mobile/edge devices	2025	MiniMax	Open-source elements; 7B parameters

2. Image Generation and Understanding Models (Text+Image)

Model Name	Primary Purposes	Supported Modalities	Release Year	Developer	Key Details
GPT-4V (GPT-4o)	VQA, captioning, image editing, reasoning	Text, Image	2023/2025	OpenAI	Closed-source; emergent abilities; high-res support
LLaVA	Instruction-tuned VQA, medical imaging	Text, Image	2023–2024	Academic	Open-source; biomedical variant available
BLIP-2	Captioning, VQA, zero-shot transfer	Text, Image	2023	Salesforce	Open-source; Q-Former for fusion

Model Name	Primary Purposes	Supported Modalities	Release Year	Developer	Key Details
MiniGPT-4	Dialogue, code gen from images, high-res	Text, Image	2023	Zhu et al.	Open-source; mobile-friendly
mPLUG-Owl	Document understanding, OCR-free parsing	Text, Image	2023	Alibaba	Open-source; modular design

3. Speech Models (TTS/STT)

Model Name	Primary Purposes	Supported Modalities	Release Year	Developer	Key Details
Whisper	STT transcription, multilingual	Audio → Text	2023–2025	OpenAI	Open-source; 99% accuracy; real-time
GPT-4o Transcribe	STT for meetings, low-latency	Audio → Text	2025	OpenAI	Closed-source; integrates with GPT
ElevenLabs	TTS voice cloning, STT	Text ↔ Audio	2025	ElevenLabs	Closed-source; emotional voices
Deepgram Nova	Real-time STT, diarization	Audio → Text	2025	Deepgram	Closed-source; low latency

4. Video Generation and Understanding Models

Model Name	Primary Purposes	Supported Modalities	Release Year	Developer	Key Details
Sora	Text-to-video, storyboarding	Text → Video	2024–2025	OpenAI	Closed-source; realistic physics
Veo 3	End-to-end video creation, editing	Text/Image → Video	2025	Google	Closed-source; high-fidelity
Runway Gen-3	Generative video, advanced editing	Text/Image → Video	2025	Runway	Closed-source; motion control
Mochi 1	Open-source video gen	Text → Video	2025	Genmo	Open-source; diffusion-based
Video-LLaMA	Video dialogue, understanding	Text, Video	2023	Zhang et al.	Open-source; EVA-CLIP for frames

5. Other Scenarios (3D, Any-to-Any Multimodal)

Model Name	Primary Purposes	Supported Modalities	Release Year	Developer	Key Details
PointLLM	3D point cloud reasoning	Text, 3D	2023	Xu et al.	Open-source; AR/VR scenarios
AnyGPT	Discrete sequence any-to-any	Text, Image, Audio, Video	2024	Not specified	Open-source; conversational
CoDi-2	Interleaved gen (multi-turn)	Text, Image, Video, Audio	2023	Tang et al.	Open-source; latent features
M2-Omni	Omni-multimodal perception	Text, Image, Video, Audio	2025	Not specified	Closed-source; comprehensive benchmarks
CAT	Audio-visual reasoning	Text, Audio, Image/Video	2024	Ye et al.	Open-source; question-aware fusion