Current AI Models in 2025: Text, Image, Video, Audio, Code

Introduction

Artificial Intelligence (AI) models have become an essential part of many areas in society and work. By 2025, AI systems can generate and understand text, create images and videos, make audio, and even write computer code. The newest AI models are used in offices, schools, entertainment, and many other fields. This essay will introduce the main AI models in 2025 for text, image, video, audio, and code. For each model, we will explain what it does, what it is best at, give simple examples, discuss how it is used in offices (like for developers, marketing, sales, and human resources), and describe its strengths and limitations. At the end, a comparison table will summarize the main differences.

Text Models

Text models are AI systems that read, write, and understand human language. They can answer questions, write emails, summarize documents, and even chat with people. In 2025, the leading text models are GPT-5, Claude 3.5, Gemini Ultra 2, Llama 3.2, and Mistral.

GPT-5

What it does: GPT-5 is a large language model from OpenAI. It can write stories, answer questions, summarize information, and have conversations.

Best at: Writing long texts, answering complicated questions, and helping with research.

Example: If you write "Explain why the sky is blue," GPT-5 gives a full answer about sunlight and the atmosphere.

Office use cases: - **Developer:** Writes code comments, helps debug errors. - **Marketing:** Writes blog posts, creates advertisements. - **Sales:** Drafts sales emails and pitches. - **HR:** Writes job descriptions and company policies.

Strengths: Very good at understanding context and creating clear, long answers.

Limitations: Sometimes makes mistakes or gives information that sounds correct but is not ("hallucinations"). Needs a lot of computer power.

Claude 3.5

What it does: Claude 3.5 is made by Anthropic. It focuses on safe, helpful, and honest answers.

Best at: Sensitive topics, ethical questions, and giving balanced advice.

Example: If asked about a difficult workplace issue, Claude 3.5 gives careful, fair advice.

Office use cases: - **Developer:** Suggests ethical code practices. - **Marketing:** Checks ad content for fairness. - **HR:** Handles sensitive employee questions.

Strengths: Safer and more careful than many other models.

Limitations: Sometimes too cautious, and may refuse to answer.

Gemini Ultra 2

What it does: Gemini Ultra 2 is from Google DeepMind. It is designed for connecting with Google's tools and handling very large documents.

Best at: Summarizing long files, finding information in Google Docs, Sheets, and Slides.

Example: Summarizes a 100-page report into a one-page summary.

Office use cases: - **Developer:** Searches through codebases. - **Marketing:** Finds trends in customer data. - **Sales:** Prepares summaries of client conversations. - **HR:** Reviews and summarizes company handbooks.

Strengths: Works well with large, complex information.

Limitations: Sometimes slower on big tasks.

Llama 3.2

What it does: Llama 3.2 is an open-source language model from Meta. It is easy to run on many computers.

Best at: Fast answers, running on smaller devices, easy customization.

Example: Used in a chatbot on a company website.

Office use cases: - **Developer:** Adds chatbots to apps. - **Marketing:** Answers common customer questions. - **Sales:** Handles simple customer support. - **HR:** Answers FAQs for employees.

Strengths: Free to use and easy to adapt.

Limitations: Not as powerful as the biggest models.

Mistral

What it does: Mistral is another open-source text model. It is made for speed and privacy.

Best at: Quick responses, running without internet, keeping data private.

Example: Used in a company’s private network to answer employee questions.

Office use cases: - **Developer:** Local code assistants. - **Marketing:** On-site content generation. - **Sales:** Offline customer support. - **HR:** Internal document search.

Strengths: Fast, private, and secure.

Limitations: Smaller knowledge base than other models.

Image Models

Image models make pictures from text or other images. The main image models in 2025 are Midjourney v7, DALL·E 4, Stable Diffusion 3, and Ideogram.

Midjourney v7

What it does: Midjourney v7 creates high-quality, artistic images from text prompts.

Best at: Art, concept design, creative projects.

Example: “Draw a futuristic city at sunset” gives a beautiful, detailed picture.

Office use cases: - **Developer:** Makes app icons or UI mockups. - **Marketing:** Designs posters, social media images. - **Sales:** Creates product mockups. - **HR:** Makes images for company newsletters.

Strengths: Very good at art styles and creativity.

Limitations: Sometimes less accurate for realistic images.

DALL·E 4

What it does: DALL·E 4 is from OpenAI. It generates realistic images from text.

Best at: Photorealistic pictures, editing images.

Example: “A cat sitting on a red sofa in a modern living room” gives a real-looking photo.

Office use cases: - **Developer:** Makes product photos for apps. - **Marketing:** Edits and creates ad images. - **Sales:** Customizes client images. - **HR:** Makes team photos for websites.

Strengths: High realism, can edit and change images.

Limitations: May struggle with very complex scenes.

Stable Diffusion 3

What it does: Stable Diffusion 3 is open-source and allows users to create images on their own computers.

Best at: Custom images, privacy, and running locally.

Example: Making custom company logos without sending data online.

Office use cases: - **Developer:** Integrates custom image generation in apps. - **Marketing:** Makes unique brand assets. - **Sales:** Creates images for private client work. - **HR:** Designs internal training materials.

Strengths: Private, flexible, and free.

Limitations: Needs setup and technical skill.

Ideogram

What it does: Ideogram is focused on making clear images with text (like posters and banners).

Best at: Adding readable text to images.

Example: "Happy Birthday" poster with beautiful fonts.

Office use cases: - **Developer:** Makes app splash screens. - **Marketing:** Designs banners and flyers. - **Sales:** Creates branded sales materials. - **HR:** Prepares event graphics.

Strengths: Good at combining text and images.

Limitations: Less good at detailed artwork.

Video Models

Video models make moving images from text or pictures. The top video models are OpenAI Sora, Runway Gen-3, and Pika Labs 2.

OpenAI Sora

What it does: Sora generates short, realistic videos from text prompts.

Best at: Creating video ads, quick explainer videos.

Example: “A dog running through a park in spring” produces a short, life-like video.

Office use cases: - **Developer:** Makes demo videos for software. - **Marketing:** Produces social media ads. - **Sales:** Sends video proposals to clients. - **HR:** Makes onboarding videos.

Strengths: Realistic, fast, easy to use.

Limitations: Limited to short videos (usually under 1 minute).

Runway Gen-3

What it does: Runway Gen-3 is a creative video generator for storytelling and editing.

Best at: Artistic videos, special effects.

Example: Turning a single photo into a moving, animated scene.

Office use cases: - **Developer:** Animates app tutorials. - **Marketing:** Makes creative brand videos. - **Sales:** Enhances product demos. - **HR:** Designs fun staff announcement videos.

Strengths: Creative, flexible, good for special effects.

Limitations: Sometimes less realistic than Sora.

Pika Labs 2

What it does: Pika Labs 2 focuses on fast video creation and editing.

Best at: Quick video edits, adding effects, simple animations.

Example: Adding text overlays or transitions to company videos.

Office use cases: - **Developer:** Edits training videos. - **Marketing:** Speeds up content creation. - **Sales:** Personalizes client videos. - **HR:** Updates staff videos quickly.

Strengths: Fast, easy for beginners.

Limitations: Not great for high-end video quality.

Audio Models

Audio models make or edit sounds, voices, and music. The newest audio models are GPT-Voice and ElevenLabs v3.

GPT-Voice

What it does: GPT-Voice can read text aloud, generate voices, and even have phone-like conversations.

Best at: Lifelike voice, many accents and styles.

Example: Reading a news article in a natural voice.

Office use cases: - **Developer:** Adds voice to apps. - **Marketing:** Makes audio ads. - **Sales:** Creates voicemail greetings. - **HR:** Reads training materials for staff.

Strengths: Very realistic, customizable voices.

Limitations: Needs careful setup for long content.

ElevenLabs v3

What it does: ElevenLabs v3 is focused on cloning voices and making custom audio.

Best at: Copying real people's voices, making audiobooks.

Example: Creating an audiobook in the CEO's voice.

Office use cases: - **Developer:** Personalizes app voices. - **Marketing:** Custom voices in ads. - **Sales:** Sends audio messages to clients. - **HR:** Makes custom onboarding audio.

Strengths: High-quality voice clones.

Limitations: Raises privacy and ethical issues if misused.

Code Models

Code models help developers write, fix, and understand programming code. The top code models are Copilot, GPT-Coder, and DeepSeek Coder.

Copilot

What it does: Copilot suggests code as you type, completes functions, and helps with debugging.

Best at: Writing code quickly, helping beginners.

Example: Auto-completing a function in Python.

Office use cases: - **Developer:** Speeds up programming tasks. - **Marketing:** Makes website changes. - **Sales:** Customizes product demos. - **HR:** Builds simple HR tools.

Strengths: Easy to use, saves time.

Limitations: Suggestions can be wrong; needs review.

GPT-Coder

What it does: GPT-Coder writes full programs from descriptions, explains code, and finds bugs.

Best at: Turning ideas into code, learning new languages.

Example: “Make a calculator app” gives a full code example.

Office use cases: - **Developer:** Rapid prototyping. - **Marketing:** Builds campaign tools. - **Sales:** Prepares demo scripts. - **HR:** Automates HR tasks.

Strengths: Very flexible, supports many languages.

Limitations: May generate insecure or inefficient code.

DeepSeek Coder

What it does: DeepSeek Coder is for advanced code search and understanding big codebases.

Best at: Finding code snippets, understanding complex projects.

Example: Searching for all login functions in a large project.

Office use cases: - **Developer:** Maintains large codebases. - **Marketing:** Audits website code. - **Sales:** Checks demo scripts. - **HR:** Finds HR software tools.

Strengths: Excellent search and analysis.

Limitations: Needs big datasets to work best.

Multimodal Models

Multimodal models handle text, images, video, and audio together. The top multimodal models are GPT-5 Omni, Gemini 2.0, and Claude 3.5 Vision.

GPT-5 Omni

What it does: GPT-5 Omni can understand and create content across text, images, and audio.

Best at: Complex tasks, like reading a picture and writing a story about it.

Example: Given a photo of a chart, writes a summary of the data.

Office use cases: - **Developer:** Builds smart apps with vision and text. - **Marketing:** Analyzes ad images and writes reports. - **Sales:** Reads client documents and creates presentations. - **HR:** Reviews resumes (text and image).

Strengths: Handles many types of data at once.

Limitations: Needs more computer power.

Gemini 2.0

What it does: Gemini 2.0 is made for Google's ecosystem and combines text, image, and video.

Best at: Searching across emails, images, and files.

Example: Finds all photos of last year's office party and writes a summary.

Office use cases: - **Developer:** Smart file search tools. - **Marketing:** Analyzes campaign data (text, images). - **Sales:** Prepares multimedia sales decks. - **HR:** Organizes staff records.

Strengths: Works well in office settings.

Limitations: Less customizable.

Claude 3.5 Vision

What it does: Claude 3.5 Vision analyzes and generates both text and images.

Best at: Explaining images, checking for bias in visuals.

Example: Reviews a flyer and checks if it is inclusive.

Office use cases: - **Developer:** Tests UI designs. - **Marketing:** Reviews ad images. - **Sales:** Checks presentation visuals. - **HR:** Examines staff photos for diversity.

Strengths: Focused on ethics and fairness.

Limitations: Limited to supported types of data.

Comparison Table

Model Type	Model Name	Strengths	Limitations	Example Use Case
Text	GPT-5	Powerful, detailed	May hallucinate	Research assistant

Model Type	Model Name	Strengths	Limitations	Example Use Case
Text	Claude 3.5	Safe, ethical	Sometimes too cautious	Workplace advice
Text	Gemini Ultra 2	Handles large files	Slower on big tasks	Summarizing reports
Text	Llama 3.2	Open-source, fast	Less powerful	Chatbots
Text	Mistral	Private, secure	Smaller knowledge base	Internal Q&A
Image	Midjourney v7	Creative, artistic	Less realistic	Concept art
Image	DALL·E 4	Realistic images	Complex scenes difficult	Ad images
Image	Stable Diffusion 3	Custom, private	Needs setup	Custom logos
Image	Ideogram	Text in images	Simple artwork	Posters and banners
Video	OpenAI Sora	Realistic short videos	Short duration	Video ads
Video	Runway Gen-3	Special effects, creative	Less realistic	Animated stories
Video	Pika Labs 2	Fast editing	Lower quality	Quick video edits
Audio	GPT-Voice	Lifelike, flexible	Set-up for long content	Voiceovers
Audio	ElevenLabs v3	Voice cloning	Privacy concerns	Audiobooks
Code	Copilot	Easy, fast	Needs review	Code completion
Code	GPT-Coder	Flexible, multi-language	May be insecure	Rapid prototyping
Code	DeepSeek Coder	Code search, analysis	Needs big datasets	Codebase analysis
Multimodal	GPT-5 Omni	Handles all data types	High resource use	Smart assistants
Multimodal	Gemini 2.0	Office integration	Less flexible	Smart search
Multimodal	Claude 3.5 Vision	Ethical, fair	Data type limits	Image review

Conclusion

AI models in 2025 are powerful tools for text, image, video, audio, and code tasks. They help in many office roles, from development to marketing and HR. Each model has strengths—like creativity, privacy, or speed—and its own limits, such as the need for review or technical setup. Understanding these models helps people and companies choose the best tools for their needs, while also being careful about privacy, ethics, and responsible use. It is important for everyone to learn how to use these AI systems wisely, as they continue to shape our workplaces and daily life (Chowdhury, 2025; Coelho, 2025; Weichert & Eldardiry, 2025; Siddharth et al., 2025; Feffer et al., 2023).

References

Chowdhury, T. (2025). Computational Thinking with Computer Vision: Developing AI Competency in an Introductory Computer Science Course. https://arxiv.org/pdf/2503.19006v1

Coelho, G. (2025). AI in Music and Sound: Pedagogical Reflections, Post-Structuralist Approaches and Creative Outcomes in Seminar Practice. https://arxiv.org/pdf/2511.17425v1

Weichert, J., & Eldardiry, H. (2025). Educating a Responsible AI Workforce: Piloting a Curricular Module on AI Policy in a Graduate Machine Learning Course. https://arxiv.org/pdf/2502.07931v1

Siddharth, S., Prince, B., Harsh, A., & Ramachandran, S. (2025). The World of AI: A Novel Approach to AI Literacy for First-year Engineering Students. https://arxiv.org/pdf/2506.08041v1

Feffer, M., Martelaro, N., & Heidari, H. (2023). The AI Incident Database as an Educational Tool to Raise Awareness of AI Harms: A Classroom Exploration of Efficacy, Limitations, & Future Improvements. https://arxiv.org/pdf/2310.06269v1