Generative artificial intelligence (Generative AI, GenAI,[1] or GAI) is a subfield of artificial intelligence that uses generative models to produce text, images, videos, or other forms of data.[2][3][4] These models learn the underlying patterns and structures of their training data and use them to produce new data[5][6] based on the input, which often comes in the form of natural language prompts.[7][8]

Generative AI tools have become more common since the AI boom in the 2020s. This boom was made possible by improvements in transformer-based deep neural networks, particularly large language models (LLMs). Major tools include chatbots such as ChatGPT, Copilot, Gemini, Claude, Grok, and DeepSeek; text-to-image models such as Stable Diffusion, Midjourney, and DALL-E; and text-to-video models such as Veo, LTXV and Sora.[9][10][11][12][13] Technology companies developing generative AI include OpenAI, Anthropic, Meta AI, Microsoft, Google, DeepSeek, and Baidu.[7][14][15]

Generative AI has raised many ethical questions and governance challenges as it can be used for cybercrime, or to deceive or manipulate people through fake news or deepfakes.[16][17] Even if used ethically, it may lead to mass replacement of human jobs.[18] The tools themselves have been criticized as violating intellectual property laws, since they are trained on copyrighted works.[19]

Generative AI is used across many industries. Examples include software development,[20] healthcare,[21] finance,[22] entertainment,[23] customer service,[24] sales and marketing,[25] art, writing,[26] fashion,[27] and product design.[28]