

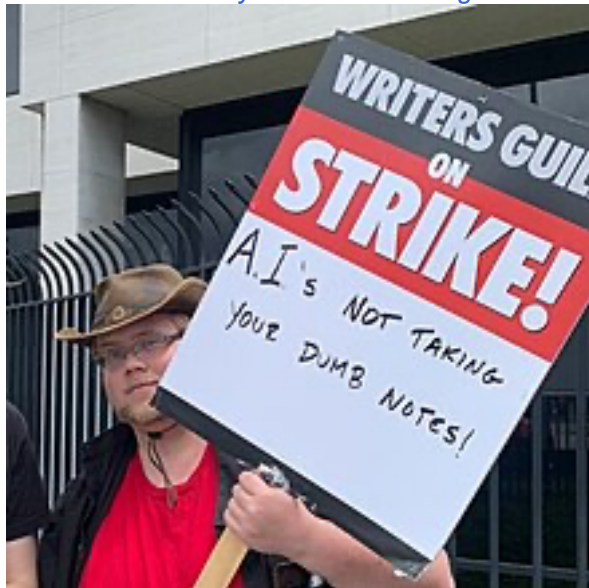
Generative artificial intelligence or **generative AI** is a type of [artificial intelligence](#) (AI) system capable of generating text, images, or other media in response to [prompts](#).^{[1][2]} Generative AI models learn the patterns and structure of their input [training data](#), and then generate new data that has similar characteristics.^{[3][4]}

Notable generative AI systems include [ChatGPT](#) (and its variant [Bing Chat](#)), a [chatbot](#) built by [OpenAI](#) using their [GPT-3](#) and [GPT-4 foundational large language models](#),^[5] and [Bard](#), a chatbot built by [Google](#) using their [LaMDA](#) foundation model.^[6] Other generative AI models include [artificial intelligence art](#) systems such as [Stable Diffusion](#), [Midjourney](#), and [DALL-E](#).^[7]

Generative AI has potential applications across a wide range of industries, including art, writing, software development, product design, healthcare, finance, gaming, marketing, and fashion.^{[8][9][10]} Investment in generative AI surged during the early 2020s, with large companies such as Microsoft, Google, and Baidu as well as numerous smaller firms developing generative AI models.^{[1][11][12]} However, there are also concerns about the potential misuse of generative AI, such as in creating [fake news](#) or [deepfakes](#), which can be used to deceive or manipulate people.^[13]

History^[edit]

Main article: [History of artificial intelligence](#)



A picketer at the [2023 Writers Guild of America strike](#). While not a top priority, one of the WGA's 2023 requests was "regulations around the use of (generative) AI".^[14]

Since its founding, the field of [machine learning](#) has used [statistical models](#), including [generative models](#), to model and predict data. Beginning in the late 2000s, the emergence of [deep learning](#) drove progress and research in image and video processing, text analysis, speech recognition, and other tasks. However, most deep neural networks were trained as [discriminative](#) models performing classification tasks such as [convolutional neural network](#)-based [image classification](#).

In 2014, advancements such as the [variational autoencoder](#) and [generative adversarial network](#) produced the first practical deep neural networks capable of learning generative, rather than discriminative, models of complex data such as images. These deep generative models were the first able to output not only class labels for images, but to output entire images.^[15]

In 2017, the [Transformer](#) network enabled advancements in generative models, leading to the first [Generative pre-trained transformer](#) in 2018.^[16] This was followed in 2019 by [GPT-2](#) which demonstrated the ability to generalize unsupervised to many different tasks as a [Foundation model](#).^[17]

In 2021, the release of [DALL-E](#), a transformer-based pixel generative model, followed by [Midjourney](#) and [Stable Diffusion](#) marked the emergence of practical high-quality [artificial intelligence art](#) from natural language prompts.

In January 2023, *Futurism.com* broke the story that [CNET](#) had been using an undisclosed internal AI tool to write at least 77 of its stories; after the news broke, CNET posted corrections to 41 of the stories.^[18]

In March 2023, [GPT-4](#) was released. A team from Microsoft Research argued that "it could reasonably be viewed as an early (yet still incomplete) version of an [artificial general intelligence](#) (AGI) system".^[19]

In April 2023, German tabloid *Die Aktuelle* published a fake AI-generated interview with reclusive former racing driver [Michael Schumacher](#). The story included two possible disclosures: the cover included the line "deceptively real", and inside the magazine acknowledged at the end of the interview that the interview was AI-generated. The editor-in-chief was fired shortly thereafter amid the controversy.^[20]

Modalities^[edit]



[Théâtre d'Opéra Spatial](#), an image generated

by [Midjourney](#)

A generative AI system is constructed by applying [unsupervised](#) or [self-supervised machine learning](#) to a data set. The capabilities of a generative AI system depend on the [modality or type](#) of the data set used.

Generative AI can be either *unimodal* or *multimodal*; unimodal systems take only one type of input, whereas multimodal systems can take more than one type of input.^[21] For example, one version of [OpenAI](#)'s [GPT-4](#) accepts both text and image inputs.^[22]

- **Text:** Generative AI systems trained on words or [word tokens](#) include [GPT-3](#), [LaMDA](#), [LLaMA](#), [BLOOM](#), [GPT-4](#), and others (see [List of large language models](#)). They are capable of [natural language processing](#), [machine translation](#), and [natural language generation](#) and can be used as [foundation models](#) for other tasks.^[23] Data sets include [BookCorpus](#), [Wikipedia](#), and others (see [List of text corpora](#)).
- **Code:** In addition to [natural language](#) text, large language models can be trained on [programming language](#) text, allowing them to generate [source code](#) for new [computer programs](#).^[24] Examples include [OpenAI Codex](#).

- **Images:** Generative AI systems trained on sets of images with [text captions](#) include [Imagen](#), [DALL-E](#), [Midjourney](#), [Adobe Firefly](#), [Stable Diffusion](#) and others (see [Artificial intelligence art](#), [Generative art](#), and [Synthetic media](#)). They are commonly used for [text-to-image](#) generation and [neural style transfer](#).^[25] Datasets include [LAION-5B](#) and others (See [Datasets in computer vision](#)).
- **Molecules:** Generative AI systems can be trained on sequences of [amino acids](#) or molecular representations such as [SMILES](#) representing DNA or proteins. These systems, such as [AlphaFold](#), are used for [protein structure prediction](#) and [drug discovery](#).^[26] Datasets include [various biological datasets](#).
- **Music:** Generative AI systems such as MusicLM can be trained on the audio waveforms of recorded music along with text annotations, in order to generate new musical samples based on text descriptions such as *a calming violin melody backed by a distorted guitar riff*.^[27]
- **Video:** Generative AI trained on annotated video can generate temporally-coherent video clips. Examples include Gen1 and Gen2 by RunwayML^[28] and Make-A-Video by [Meta Platforms](#).^[29]
- **Robot actions:** Generative AI trained on the motions of a robotic system can generate new trajectories for [motion planning](#). For example, UniPi from Google Research uses prompts like *"pick up blue bowl"* or *"wipe plate with yellow sponge"* to control movements of a robot arm.^[30]