# What is artificial intelligence?

Excerpts from an article in The Washington Post

Artificial intelligence is an umbrella term for a vast array of technology. There is no single definition, and even researchers disagree. Generally, AI is a field of computer science that focuses on creating and training machines to perform intelligent tasks, something that, if a person was doing it, we would call it intelligence.

For decades, AI has largely been used for analysis, allowing people to spot patterns and make predictions by assessing huge sets of data.

But advancements in the field have led to a boom in generative AI, a form of artificial intelligence that can make things. The technology can create words, sounds, images and video, sometimes at a level of sophistication that mimics human creativity. It backs chatbots like ChatGPT and image generators like DALL-E.

Although this technology can't "think" like humans do, it can sometimes create work of a similar quality. AI-powered image generators have made photos that tricked art judges into thinking they were human-made, and voice generating software has preserved voices of people suffering from degenerative diseases such as ALS.

Chatbots backed by generative AI have dazzled users by carrying on eerily lifelike conversations — an early dream of the field as envisioned by Alan Turing. In 1950, he developed the "Turing test," which judged the success of an AI machine by how well it could fool users into believing it was human.

Turing never gave much credence to the idea that a computer could really "think" — he called that question "too meaningless to deserve discussion."

#### Artificial intelligence software is nothing without data.

The tools develop intelligence through machine learning, a process that allows computers to "learn" on their own, without requiring a programmer to tell them each step. Feed a computer massive amounts of data, and it eventually can recognize patterns and predict outcomes.

Key to this process are mathematical systems that act like a computerized brain, helping the technology find connections in data. They're modeled after the human brain, with layers of artificial "neurons" that communicate information to one another. Even experts don't necessarily understand all the intricacies of how these neural networks work.

Large language models learn to "speak" by collecting massive amounts of text, often websites scraped from the internet, and finding statistical relationships between words. When these systems pattern-match, it can lead to feats of creativity: A chatbot can create song lyrics closely matching Jay-Z's style because it's absorbed the patterns of his entire discography. But AI's don't have awareness of the meanings behind words.

Parameters, which are numerical points across a large language model's training data, dictate how proficient it is at its tasks, such as predicting the next word in a sentence.

In the future, some researchers say, the technology will approach artificial general intelligence, or AGI, a point at which it matches or exceeds the intelligence of humans. The idea is core to the mission of some artificial intelligence labs, like OpenAI, which lists achieving AGI as its goal in its founding documents. Other experts contest that AI is anywhere close to achieving that kind of sophistication, with some critics contending that it's a marketing term.

## How do we interact with AI? Chatbots, like ChatGPT, Bard and more.

The most common way people experience artificial intelligence is through chatbots, which work like an advanced form of instant messenger, answering questions and formulating tasks from prompts.

These bots are trained on massive collections of internet data, including Reddit conversations and digital books. Chatbots are incredibly adept at finding patterns and imitating speech, but they don't interpret meanings. It's a super, super high-fidelity version of autocomplete.

Since it debuted in November, ChatGPT has stunned users with its ability to produce fluid language — generate complete novels, computer code, TV episodes and songs. GPT stands for "generative pre-trained transformer." "Generative," meaning that it uses AI to create things. "Pre-trained," means that it has already been trained on a large amount of data. And "transformer" is a powerful type of neural network that can process language.

Created by the San Francisco start-up OpenAI, ChatGPT has led to a rush of companies releasing their own chatbots. Microsoft's chatbot, Bing, uses the same underlying technology as ChatGPT. And Google released a chatbot, Bard, based on the company's LaMDA model.

Some people think chatbots will alter how people find and consume information on the internet. Instead of entering a term into a search engine, like Google, and sifting through various links, people may end up asking a chatbot a question and getting a confident answer back. (Though sometimes these answers are false — stay tuned!)

## Taming AI: Deepfakes, hallucination and misinformation

The boom in generative artificial intelligence brings exciting possibilities — but also concerns that the cutting-edge technology might cause harm.

Chatbots can sometimes make up sources or confidently spread misinformation. In one instance, ChatGPT invented a sexual harassment scandal against a college law professor. It can also churn out conspiracy theories and racist answers. Sometimes it expresses biases in its work: In one experiment, robots identified Black men when asked to find a "criminal" and marked all "homemakers" as women.

AI ethicists and researchers have long been concerned that, because chatbots draw on massive amounts of human speech — using data from Twitter to Wikipedia — they absorb our problems and biases. Companies have tried to put semantic guardrails in place to limit what chatbots can say, but that doesn't always work.

Sometimes artificial intelligence produces information that sounds plausible but is irrelevant, nonsensical or entirely false. These odd detours are called hallucinations. Other people have become so immersed in chatbots they falsely believe the software can think, feel, and act outside of human control. Experts say it can't — at least not yet — but it can speak in a fluid way so that it mimics something alive.

Another worry is deepfakes, which are synthetically generated photos, audio or video that are fake but look real. The same technology that can produce awesome images could be deputized to fake wars, make celebrities say things they didn't actually say or cause mass confusion or harm.

Companies test their artificial intelligence models for vulnerabilities, rooting out biases and weaknesses by simulating flaws in a process called red teaming.

Despite attempts to tame the technology, the innovation and sophistication of generative AI causes some to worry.

When things talk to us like humans, we pick up a little suspension of belief. We assume that these things are trying to be faithful to us, and when they come across as

authoritative, we can find it hard to be skeptical. (How many times have you argued with Amazon's Alexa?)

#### **Companies in the chatbot race**

OpenAI: The San Francisco-based artificial intelligence research lab launched as a nonprofit to build "artificial general intelligence" outside of Big Tech's control. Since then, it's transformed into a major corporate player, creating image generator DALL-E and chatbot ChatGPT. It is now for-profit and has partnered with companies including Microsoft and Salesforce.

Google: The tech giant — long a leader in AI including via search — launched chatbot Bard after competitors' offerings went viral. It is known for its LaMDA technology, a system for building chatbots based on large language models.

Microsoft: The software company invested billions of dollars in OpenAI and teamed up to create a Bing chatbot, developed on GPT-4 technology. But there have been missteps, including when the chatbot went rogue, told reporters it has feelings and called itself Sydney — forcing the tech giant to reel it back in some ways.

Meta: Even before ChatGPT, Facebook's parent company released a chatbot called Blenderbot, but it failed to gain traction. Its chief artificial intelligence scientist later called the bot "boring" because it was "made safe."

IBM: IBM was an early leader in artificial intelligence close to the current chatbot trends, most notably with its robot Watson, which captivated audiences on "Jeopardy!"