The history of artificial intelligence

Humans have dreamed of creating thinking machines from ancient times. Folklore and historical attempts to build programmable devices reflect this long-standing ambition and fiction abounds with the possibilities of intelligent machines, imagining their benefits and dangers. It's no wonder that when OpenAI released the first version of GPT (Generative Pretrained Transformer), it quickly gained widespread attention, marking a significant step toward realizing this ancient dream.

GPT-3 was a landmark moment in AI due to its unprecedented size, featuring 175 billion parameters, which enabled it to perform a wide range of natural language tasks without extensive fine-tuning. This model was trained using big data, allowing it to generate human-like text and engage in conversations. It also had the ability to perform few-shot learning, significantly improving its versatility and demonstrated usefulness in commercial AI applications such as chatbots and virtual assistants.

Today, AI is increasingly becoming embedded into many aspects of daily life, from social media to work processes and as the technology improves, its influence will continue to grow. To understand the directions the technology can take, it helps to understand how we got here. Here is a history of major developments in AI:

ⓘ

Industry newsletter

The latest AI trends, brought to you by experts

Get curated insights on the most important—and intriguing—AI news. Subscribe to our weekly Think newsletter. See the IBM Privacy Statement.

johndoe@yourdomain.com   Subscribe

Pre-20th century

1726

Jonathan Swift's fantastic novel "Gulliver's Travels" introduces the idea of The Engine, a large mechanical contraption used to assist scholars in generating new ideas, sentences and books.

Scholars turn handles on the machine, which rotates wooden blocks inscribed with words. The machine is said to create new ideas and philosophical treatises by combining words in different arrangements:

> "Every one knew how laborious the usual method is of attaining to arts and sciences; whereas by his contrivance the most ignorant person, at a reasonable charge and with a little bodily labour, might write books in philosophy, poetry, politics, laws, mathematics and theology, without the least assistance from genius or study."
>
> - Jonathan Swift's Gulliver's Travels (1726)

Swift's satire anticipates the concept of algorithmic text generation, which is now a reality with modern AI. AI models can produce coherent text by combining words and ideas based on underlying algorithms, similar to what Swift's fictional Engine is meant to do.

AI Academy

Become an AI expert

Gain the knowledge to prioritize AI investments that drive business growth.
Get started with our free AI Academy today and lead the future of AI in your

organization.

1900–1950

1914

Spanish engineer Leonardo Torres y Quevedo demonstrates the first chess-playing machine, *El Ajedrecista* at the *Exposition Universelle* in Paris. It used electromagnets and was fully automated. *El Ajedrecista* automatically played a simple chess endgame of king and rook versus king. The machine required no human intervention once it was set up—it autonomously made legal chess moves and if the human opponent made an illegal move, the machine would signal the error. If the machine was placed in a winning position, it was able to checkmate the human opponent reliably.

1921

A play named "Rossum's Universal Robots" (R.U.R) opens in London. The play by Karel Čapek is the first time the word "robot" is used in English. In Czech, the word "robota" is associated with compulsory or forced work performed by peasants in a feudal system. The term "robot" quickly gained international recognition after the play's success and became the standard term for mechanical or artificial beings created to perform tasks. Though Čapek's robots are organic, the word came to be associated with mechanical, humanoid machines designed to perform monotonous, unskilled labor.

1939

John Vincent Atanasoff, a professor of physics and mathematics at Iowa State College, and his graduate student Clifford Berry, create the Atanasoff-Berry Computer (ABC) with a grant of USD 650 at Iowa State University. The ABC computer is considered one of the earliest digital electronic computers and a milestone in the field of American computer science.

While the ABC is never fully operational or widely used, it introduced several key concepts that would become foundational in the development of modern computing.

Unlike previous computing devices that relied on decimal systems, the ABC used binary (1s and 0s) to represent data, which became the standard for computers thereafter. ABC is also one of the first computers to use electronic circuits for computation instead of mechanical or electromechanical systems, allowing for faster and more reliable calculations. The ABC

separated data storage (memory) from the processing unit (logic operations), a principle still followed in modern computer architecture. It used capacitors to store data and could handle up to 30 simultaneous equations.

ABC employed around 300 vacuum tubes for its logic operations, which made it much faster than earlier mechanical calculators. Vacuum tubes, though bulky and prone to failure, are a key development in electronic computing. The ABC weighed over 700 pounds and could solve up to 29 simultaneous linear equations.

1943

Warren S. McCulloch and Walter Pitts publish "A Logical Calculus of the Ideas Immanent in Nervous Activity" in the *Bulletin of Mathematical Biophysics*.[1] It is one of the seminal works in the history of both neuroscience and AI. The paper lays the foundation for the idea that the brain can be understood as a computational system and it introduces the concept of artificial neural networks, now a key technology in modern AI. This idea inspires computer systems that simulate brain-like functions and processes, particularly through neural networks and deep learning.

1950

British mathematician Alan Turing's landmark paper "Computing Machinery and Intelligence" is published in *Mind*.[2] This paper is a foundational text in AI and addresses the question, "Can machines think?" Turing's approach established a foundation for future discussions on the nature of thinking machines and how their intelligence might be measured via the "imitation game," now known as the Turing Test. Turing introduced a thought experiment to avoid directly answering the question "Can machines think?" Instead, he rephrased the problem into a more specific, operational form: Can a machine exhibit intelligent behavior indistinguishable from that of a human?

The Turing Test has become a central concept in AI, serving as one way to measure machine intelligence by assessing a machine's ability to convincingly mimic human conversation and behavior.

1950–1980

1951

Marvin Minsky and Dean Edmunds build the first artificial neural network. The Stochastic Neural Analog Reinforcement Calculator (SNARC) is an early attempt to model learning processes in the human brain, specifically through reinforcement learning.

SNARC is designed to simulate the behavior of a rat navigating a maze. The idea is to have the machine mimic the way animals learn through rewards and punishment—adjusting its behavior over time based on feedback. It is an analog computer using a network of 3000 vacuum tubes alongside synaptic weights to simulate 40 neuron-like units.

1952

Allen Newell, a mathematician and computer scientist, and Herbert A. Simon, a political scientist, develop influential programs such as the Logic Theorist and General Problem Solver, which are among the first to mimic human problem-solving abilities using computational methods.

1955

The term "artificial intelligence" is first coined in a workshop proposal titled "A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence,"[3] submitted by John McCarthy of Dartmouth College, Marvin Minsky of Harvard University, Nathaniel Rochester from IBM and Claude Shannon from Bell Telephone Laboratories.

The workshop, which took place a year later, in July and August 1956, is generally considered the official birthdate of the burgeoning field of AI.

1957

Frank Rosenblatt, a psychologist and computer scientist, develops the Perceptron, an early artificial neural network that enables pattern recognition based on a two-layer computer learning network. The Perceptron introduces the concept of a binary classifier that can learn from data by adjusting the weights of its inputs through learning algorithms. While limited to solving linearly separable problems, it laid the foundation for future neural networks and machine learning developments.

1958

John McCarthy develops programming language Lisp[4], which stands for LISt Processing. Lisp is developed out of McCarthy's work on formalizing algorithms and mathematical logic, particularly influenced by his desire to create a programming language that can handle symbolic information. Lisp soon becomes the most popular programming language used in AI research.

1959

Arthur Samuel pioneers the concept of machine learning by developing a computer program that improves its performance at checkers over time. Samuel demonstrates that a computer can be programmed to follow predefined rules and "learn" from experience, eventually playing better than the programmer. His work marks a major step toward teaching machines to improve through experience, coining the term "machine learning" in the process.

Oliver Selfridge publishes his paper "Pandemonium: A paradigm for learning."[5] His pandemonium model proposed a system in which various "demons" (processing units) work together to recognize patterns. The demons compete to identify features in data that has not been preprogrammed, simulating unsupervised learning. Selfridge's model is an early contribution to pattern recognition, influencing future developments in machine vision and AI.

John McCarthy introduces the concept of the Advice Taker in his paper "Programs with Common Sense."[6] This program aims to solve problems by manipulating sentences in formal logic, laying the groundwork for reasoning in AI. McCarthy envisions a system that can understand instructions, reason with common-sense knowledge and learn from experience, with the long-term goal of developing AI that can adapt and learn as effectively as humans. This concept helps shape early research in knowledge representation and automated reasoning.

1965

Philosopher Hubert Dreyfus publishes "Alchemy and Artificial Intelligence,"[7] arguing that the human mind operates fundamentally differently from computers. He predicts limits to AI progress due to the challenges of replicating human intuition and understanding. His critique is influential in sparking debates about AI's philosophical and practical limits.

I.J. Good writes "Speculations Concerning the First Ultraintelligent Machine,"[8] famously asserting that once an ultraintelligent machine is created, it can design even more intelligent systems, making it humanity's last invention—provided it remains controllable. His ideas prefigure modern discussions on AI superintelligence and its risks.

Joseph Weizenbaum develops ELIZA,[9] a program that mimics human conversation by responding to typed input in natural language. Although Weizenbaum intends to show the

superficiality of human-computer communication, he is surprised by how many users attributed human-like emotions to the program, raising ethical questions about AI and human interaction.

Edward Feigenbaum, Bruce Buchanan, Joshua Lederberg and Carl Djerassi developed DENDRAL at Stanford University.[10] It is the first expert system to automate the decision-making process of organic chemists by simulating hypothesis formation. DENDRAL's success marks an advance in AI, demonstrating how systems can perform specialized tasks as well as or better than human experts.

1966

Developed at SRI in the late 1960s, Shakey is the first mobile robot capable of reasoning about its own actions, combining perception, planning and problem-solving.[11] In a 1970 Life magazine article, Marvin Minsky predicts that within three to eight years, AI would achieve the general intelligence of an average human. Shakey's achievements mark a milestone in robotics and AI, though Minsky's ambitious timeline proves overly optimistic.

1969

Arthur Bryson and Yu-Chi Ho introduce backpropagation, a method for optimizing multi-stage dynamic systems. While originally developed for control systems, this algorithm becomes crucial for training multilayer neural networks. Backpropagation only gained prominence in the 2000s and 2010s with advances in computing power, enabling the rise of deep learning.

Marvin Minsky and Seymour Papert publish *Perceptrons: An Introduction to Computational Geometry*,[12] which critically analyzed the limitations of single-layer neural networks. Their work is often blamed for reducing interest in neural networks. In the 1988 edition, they argue that progress had already stalled due to a lack of theoretical understanding despite numerous experiments with perceptrons by the mid-1960s.

1970

Terry Winograd creates SHRDLU, a groundbreaking natural language understanding program.[13] SHRDLU can interact with users in plain English to manipulate objects in a virtual block world, demonstrating the potential for computers to understand and respond to complex instructions. It is an early achievement in natural language processing, though its success is

limited to specific, highly structured environments. SHRDLU's capabilities highlight both the promise and the challenges of achieving broader AI language understanding.

1972

Developed at Stanford University, MYCIN is one of the first expert systems created to assist doctors in diagnosing bacterial infections and recommending antibiotic treatments.[14] MYCIN uses a rule-based approach to simulate the decision-making process of human experts and creates a platform for the development of medical AI systems. However, due to ethical and legal concerns, it is never implemented in clinical practice.

1973

James Lighthill presents a critical report to the British Science Research Council on the progress of AI research, concluding that AI has failed to deliver on its early promises.[15] He argues that the field has not produced significant breakthroughs, leading to a drastic reduction in government funding for AI in the UK. This report contributed to the onset of the first AI winter[16], a period of diminished interest and investment in AI research.

1980–2000

1980

WABOT-2[17,] a humanoid robot developed at Waseda University in Japan, is built starting in 1980 and completed around 1984. It followed WABOT-1, which had been built in 1973. While WABOT-1 focused on basic mobility and communication, WABOT-2 is more specialized, designed specifically as a musician robot. It can read musical scores with its camera "eyes," converse with humans, play music on an electronic organ and even accompany a human singer. This project represents a meaningful step toward the development of humanoid robots and AI capable of performing complex, human-like tasks such as artistic expression.

1982

Japan launched the Fifth Generation Computer Systems Project (FGCS) with the goal of developing computers that could handle logical reasoning and problem-solving, pushing AI research forward. This ambitious project aimed to build machines capable of performing tasks such as natural language processing and expert systems. Though it was halted in 1992, the FGCS project and its findings contributed greatly to the development of the concurrent logic programming field.

1984

At the annual meeting of the Association for the Advancement of Artificial Intelligence (AAAI), Roger Schank and Marvin Minsky caution about an impending "AI Winter," predicting that inflated expectations surrounding AI will soon lead to a collapse in investment and research, similar to the funding reduction in the mid-1970s. Their prediction came true within three years as interest in AI dwindled due to unmet promises, resulting in decreased funding and a slowdown in progress. This period became known as the second AI Winter.

Schank and Minsky's warning highlights the cyclical nature of AI hype, where bursts of optimism are followed by disillusionment when the technology failed to meet investors' and the public's expectations.

1986

David Rumelhart, Geoffrey Hinton and Ronald Williams publish the seminal paper "Learning representations by back-propagating errors," in which they described the backpropagation algorithm.[18] This method allows neural networks to adjust their internal weights by "back-propagating" the error through the network, improving the ability of multilayer networks to learn complex patterns. The backpropagation algorithm becomes a foundation for modern deep learning, sparking renewed interest in neural networks and overcoming some limitations highlighted in earlier AI research. This discovery builds on the 1969 work of Arthur Bryson and Yu-Chi Ho by applying the backpropagation algorithm specifically to neural networks, overcoming previous limitations in training multilayer networks.

This breakthrough makes artificial neural networks viable for practical applications and opened the door for the deep learning revolution of the 2000s and 2010s.

1987

During his Educom keynote speech, Apple CEO John Sculley presents the Knowledge Navigator video, which imagines a future where digital smart agents help users access vast amounts of information over networked systems.[19] This visionary concept depicts a professor interacting with a knowledgeable, voice-activated assistant who can retrieve data, answer questions and display information from what we now recognize as the internet. The video foresaw many elements of modern technologies such as AI assistants, networked knowledge databases and our interconnected digital world.

1988

Judea Pearl publishes *Probabilistic Reasoning in Intelligent Systems*, revolutionizing how AI processes information under uncertainty.[20] This work introduces Bayesian networks, a formalism for representing complex probability models and the algorithms for performing inference within them. Pearl's methods allowed AI systems to make reasoned decisions in uncertain environments, influencing fields far beyond AI, including engineering and the natural sciences. His contributions are recognized with his 2011 Turing Award, which cited his role in creating the "representational and computational foundation" for modern probabilistic reasoning in AI.[21]

Rollo Carpenter developed Jabberwacky[22], an early chatbot designed to simulate human-like conversations that are interesting, entertaining and humorous. Unlike rule-based systems, Jabberwacky learns from human interactions to generate more natural dialogue, paving the way for later conversational AI models. This chatbot is one of the first attempts to create AI that mimics spontaneous, everyday human conversation through continuous learning from its interactions with users.

Researchers from the IBM T.J. Watson Research Center publish "A Statistical Approach to Language Translation," marking a pivotal shift from rule-based to probabilistic methods in machine translation.[23] This approach, exemplified by IBM's Candide project24, uses 2.2 million English-French sentence pairs, primarily sourced from the Canadian Parliament's proceedings. This new methodology emphasizes learning from statistical patterns in data rather than attempting to comprehend or "understand" the languages, reflecting the broader trend toward machine learning that relies on analyzing known examples. This probabilistic model paved the way for many future advancements in natural language processing and machine translation.

Marvin Minsky and Seymour Papert release an expanded edition of their 1969 book *Perceptrons*, a seminal critique of early neural networks. In the new prologue, titled "A View from 1988," they reflected on the slow progress in the field of AI, noting that many researchers continued to repeat mistakes from the past due to unfamiliarity with earlier challenges.[12] They highlight the need for deeper theoretical understanding, which is lacking in earlier neural network research. They underscore their original criticisms while acknowledging emerging approaches that would later lead to modern deep learning advancements.

1989

Yann LeCun and a team of researchers at AT&T Bell Labs achieve a breakthrough by successfully applying the backpropagation algorithm to a multilayer neural network to recognize

handwritten ZIP code images.[24] This is one of the first practical applications of deep learning using convolutional neural networks. Despite the limited hardware of the time, it takes about three days to train the network, a meaningful improvement over earlier attempts. The system's success in handwritten digit recognition, a key task for automating postal services, demonstrates the potential of neural networks for image recognition tasks and laid the foundation for the explosive growth of deep learning in the following decades.

1993

Science fiction author and mathematician Vernor Vinge publishes the essay "The Coming Technological Singularity," in which he predicts that superhuman intelligence will be created within the next 30 years, fundamentally transforming human civilization.[25] Vinge argues that technological advances, particularly in AI, will lead to an intelligence explosion—machines surpassing human intelligence—and the end of the human era as we know it. His essay is instrumental in popularizing the concept of the "technological singularity," a moment when AI would surpass human control, sparking debate in AI, ethics and futurism communities.

This prediction continues to influence discussions about the potential impacts of AI and superintelligence, particularly the existential risks and the ethical considerations of creating machines with intelligence far beyond human capability.

1995

Richard Wallace develops the chatbot A.L.I.C.E.[26] (Artificial Linguistic Internet Computer Entity), building on the foundation laid by Joseph Weizenbaum's ELIZA program. Unlike ELIZA, which relied on scripted responses to simulate conversation, A.L.I.C.E. leveraged the newly emerging World Wide Web to collect and process vast amounts of natural language data, enabling it to engage in more complex and fluid conversations. A.L.I.C.E. uses a pattern-matching technique called AIML (Artificial Intelligence Markup Language) to parse and generate responses, making it more adaptable and scalable than its predecessors. Wallace's work sets the stage for further advancements in conversational AI, influencing modern virtual assistants and chatbots.

1997

Sepp Hochreiter and Jürgen Schmidhuber introduce Long Short-Term Memory (LSTM), a type of recurrent neural network (RNN) designed to overcome the limitations of traditional RNNs, particularly their inability to capture long-term dependencies in data effectively. LSTM networks are widely used in applications such as handwriting recognition, speech recognition, natural language processing and time series forecasting.

IBM's Deep Blue makes history by defeating reigning world chess champion Garry Kasparov in a six-game match.[27] This is the first time a computer chess-playing program beat a world champion under standard chess tournament time controls. Deep Blue's victory demonstrated that computers can outperform humans in highly strategic games, long considered a hallmark of human intelligence. The machine's ability to calculate millions of moves per second, combined with advancements in game theory and heuristics, enable it to outmaneuver Kasparov, solidifying Deep Blue's place in AI history.

The event also sparked debates about the future relationship between human cognition and AI, influencing subsequent AI research in other fields such as natural language processing and autonomous systems.

1998

Dave Hampton and Caleb Chung create Furby, the first widely successful domestic robotic pet.[28] Furby can respond to touch, sound and light and "learn" language over time, starting with its language, Furbish, but gradually "speaking" more English as it interacts with users. Its ability to mimic learning and engage with users makes it a precursor to more sophisticated social robots, blending robotics with entertainment for the first time in a consumer product.

Yann LeCun, Yoshua Bengio and their collaborators publish influential papers on neural networks' application to handwriting recognition.[29] Their work focuses on using convolutional neural networks to optimize the backpropagation algorithm, making it more effective for training deep networks. By refining the backpropagation process and demonstrating the power of CNNs for image and pattern recognition, LeCun and Bengio's research set the stage for modern deep learning techniques used in a wide range of AI applications today.

2000–2020

2000

Cynthia Breazeal at MIT develop Kismet, a robot designed to interact with human beings through emotional and social cues.[30] Kismet is equipped with cameras, microphones and expressive facial features, allowing it to perceive and respond to human emotions such as happiness, sadness and surprise. This development marks an advance in social robotics, exploring how robots can interact with humans more naturally.

2006

Geoffrey Hinton publishes "Learning Multiple Layers of Representation," which summarizes key breakthroughs in deep learning and outlines how multilayer neural networks can be trained more effectively.[31] Hinton's work focuses on training networks with graduated connections to generate sensory data rather than simply classifying it. This approach represents a shift from traditional neural networks to what we now call deep learning, allowing machines to learn complex hierarchical representations of data.

2007

Fei-Fei Li and her team at Princeton University initiate the ImageNet project, creating one of the largest and most comprehensive databases of annotated images.[32] ImageNet is designed to support the development of visual object recognition software by providing millions of labeled images across thousands of categories. The scale and quality of the dataset enables advancements in computer vision research, particularly in training deep learning models to recognize and classify objects in images.

2009

Rajat Raina, Anand Madhavan and Andrew Ng publish "Large-scale Deep Unsupervised Learning using Graphics Processors," arguing that graphics processing units (GPUs) can far outperform traditional multi-core CPUs for deep learning tasks.[33] They demonstrate that GPUs' superior computational power can revolutionize the applicability of deep unsupervised learning methods, allowing researchers to train more extensive and complex models more efficiently. This work is instrumental in accelerating the adoption of GPUs in deep learning, leading to the breakthroughs in the 2010s that power modern AI applications in fields such as computer vision and natural language processing.

Computer scientists at Northwestern University's Intelligent Information Laboratory develop Stats Monkey, a program capable of automatically generating sports news stories without human intervention.[34] Using game statistics, Stats Monkey can craft coherent narratives about baseball games, complete with recaps, player performances and analysis.

2011

IBM's Watson, an advanced natural language question-answering computer, makes headlines by competing on the game show Jeopardy! against two of the show's most successful champions, Ken Jennings and Brad Rutter, and defeating them.[35] Watson's ability to process and interpret natural language and its vast knowledge base allow it to answer complex questions quickly and accurately. This victory highlights the advancements in AI's ability to understand and interact with human language on a sophisticated level.

Apple launches Siri, a virtual assistant integrated into the iOS operating system. Siri features a natural language user interface that allows users to interact with their devices through voice commands. Siri can perform tasks such as sending messages, setting reminders, providing recommendations and answering questions using machine learning to adapt to each user's preferences and voice patterns. This personalized, adaptive voice recognition system gives users an individualized experience and marks a leap in the usability and accessibility of AI-powered assistants for everyday consumers.

2012

Jeff Dean and Andrew Ng conduct an experiment using a massive neural network with 10 million unlabeled images sourced from YouTube videos.[36] During the experiment, the network, without prior labeling, learns to recognize patterns in the data and "to our amusement," one neuron becomes particularly responsive to images of cats. This discovery is a demonstration of unsupervised learning—showing how deep neural networks can autonomously learn features from vast amounts of data.

Researchers from the University of Toronto, led by Geoffrey Hinton, design a convolutional neural network that achieves a breakthrough result in the ImageNet Large Scale Visual Recognition Challenge.[37] Their CNN, known as AlexNet achieves a 16% error rate, a substantial improvement over the previous year's best result of 25%. This achievement marks a turning point for deep learning in computer vision, proving that CNNs can outperform traditional image classification methods when trained on large datasets.

2016

Google DeepMind's AlphaGo defeated Lee Sedol, one of the world's top Go players. Go, a complex board game with more possible moves than atoms in the universe, had long been considered a challenge for AI.[38] AlphaGo's 4–1 victory over Sedol is a groundbreaking moment

in AI, showcasing the power of deep learning techniques to handle highly complex strategic tasks that had previously been beyond AI's capabilities.

Hanson Robotics introduced Sophia, a highly advanced humanoid robot.[39] Sophia can recognize faces, make eye contact and hold conversations using a combination of image recognition and natural language processing.

2017

Researchers at the Facebook Artificial Intelligence Research (FAIR) lab train two chatbots to negotiate with each other. While the chatbots are programmed to communicate in English, during their conversations, they began to diverge from the structured human language and create their own shorthand language to communicate more efficiently.[40] This development is unexpected, as the bots optimize their communication without human intervention. The experiment is halted to keep the bots within human-understandable language, but the occurrence highlights the potential of AI systems to evolve autonomously and unpredictably.

2020

OpenAI introduces GPT-3, a language model with 175 billion parameters, making it one of the largest and most sophisticated AI models to date. GPT-3 demonstrates the ability to generate human-like text, engage in conversations, write code, translate languages and generate creative writing based on natural language prompts. As one of the earliest examples of a large language model (LLM), GPTs massive size and scale enabled it to perform a wide variety of language tasks with little to no task-specific training. This example demonstrated the potential of AI to understand and produce highly coherent language.

DeepMind's AlphaFold 2 makes a breakthrough in biology by accurately predicting the 3D structures of proteins from their amino acid sequences. This achievement solves a problem that stumped scientists for decades—understanding how proteins fold into their unique three-dimensional shapes. AlphaFold 2's high accuracy in protein structure prediction has implications for disease research and drug development, offering new ways to understand the molecular mechanisms behind illnesses and to design novel therapeutics more efficiently.

2021 - Present

2021

MUM (Multitask Unified Model), developed by Google, is a powerful AI model designed to improve the search experience by understanding and generating language across[75] languages. MUM can multitask, analyzing text, images and videos simultaneously, allowing it to tackle more complex and nuanced search queries.[41] Unlike traditional models, MUM can handle multimodal inputs and provide comprehensive, context-rich answers to sophisticated questions involving multiple sources of information.

Tesla launches the Full Self-Driving (FSD) Beta, an advanced driver assistance system aimed at achieving fully autonomous driving. The FSD Beta leverages deep learning and neural networks to navigate complex driving scenarios such as city streets in real-time, highways and intersections. It allows Tesla vehicles to steer, accelerate and brake autonomously under specific conditions while requiring driver supervision. Tesla's FSD Beta marks a step toward the company's goal of fully autonomous vehicles, though regulatory challenges and safety concerns remain in the path toward achieving widespread deployment of autonomous driving technology.

2021–2023

OpenAI launches DALL-E, followed by DALL-E 2 and DALL-E 3, generative AI models capable of generating highly detailed images from textual descriptions. These models use advanced deep learning and transformer architecture to create complex, realistic and artistic images based on user input. DALL-E 2 and 3 expand the use of AI in visual content creation, allowing users to turn ideas into imagery without traditional graphic design skills.

2024

February

Google launches Gemini 1.5 in limited beta, an advanced language model capable of handling context lengths of up to 1 million tokens.[42] The model can process and understand vast amounts of information in a single prompt, improving its ability to maintain context in complex conversations and tasks over extended text. Gemini 1.5 represents a notable leap in natural language processing by providing enhanced memory capabilities and contextual understanding over long inputs.

OpenAI publicly announces Sora, a text-to-video model capable of generating videos up to one minute long from textual descriptions.[43] This innovation expands the use of AI-generated content

beyond static images, enabling users to create dynamic, detailed video clips based on prompts. Sora is expected to open new possibilities in video content creation.

StabilityAI announces Stable Diffusion 3, its latest text-to-image model. Like Sora, Stable Diffusion 3 uses a similar architecture for generating detailed and creative content from text prompts.[44]

May

Google DeepMind unveils a new extension of AlphaFold that helps identify cancer and genetic diseases, offering a powerful tool for genetic diagnostics and personalized medicine.[45]

IBM introduces the Granite™ family of generative AI models as part of its watsonx™ portfolio of AI products. Ranging 3–34 billion parameters, Granite models are designed for tasks such as code generation, time-series forecasting and document processing. Open-sourced and available under the Apache 2.0 license, these models are lightweight, cost-effective and customizable, making them ideal for a wide range of business applications.

June

Apple announces Apple Intelligence, an integration of ChatGPT into new iPhones and Siri.[46] This integration allows Siri to perform more complex tasks, hold more natural conversations and better understand and execute nuanced commands.

September

NotebookLM introduces DeepDive, a new multimodal AI capable of transforming source materials into engaging audio presentations structured as a podcast.[47] DeepDive's ability to analyze and summarize information from different formats, including webpages, text, audio and video, opens new opportunities for creating personalized and automated content across various platforms. This capability makes it a versatile tool for media production and education.

Current AI trends point to new evolutions of generative AI operating on smaller, more efficient foundation models and the rise of agentic AI, where specific AI models work together to complete user requests faster. Further into the future, autonomous vehicles will be cruising the highways, multimodal AI will create audio, video, text and images in a single platform and AI assistants will help users navigate their personal lives and careers.