

"AI" redirects here. For other uses, see [AI \(disambiguation\)](#) and [Artificial intelligence \(disambiguation\)](#).

Part of a series on

Artificial intelligence (AI)



show

[Major goals](#)

show

Approaches

show

[Applications](#)

show

[Philosophy](#)

show

[History](#)

show

Glossary

- [v](#)
- [t](#)
- [e](#)

Artificial intelligence (AI), in its broadest sense, is [intelligence](#) exhibited by [machines](#), particularly [computer systems](#). It is a [field of research](#) in [computer science](#) that develops and studies methods and [software](#) that enable machines to [perceive their environment](#) and use [learning](#) and intelligence to take actions that maximize their chances of achieving defined goals.^{[\[1\]](#)} Such machines may be called AIs.

High-profile [applications of AI](#) include advanced [web search engines](#) (e.g., [Google Search](#)); [recommendation systems](#) (used by [YouTube](#), [Amazon](#), and [Netflix](#)); [virtual assistants](#) (e.g., [Google Assistant](#), [Siri](#), and [Alexa](#)); [autonomous vehicles](#) (e.g., [Waymo](#)); [generative](#) and [creative](#) tools (e.g., [ChatGPT](#) and [AI art](#));

and [superhuman](#) play and analysis in [strategy games](#) (e.g., [chess](#) and [Go](#)). However, many AI applications are not perceived as AI: "A lot of cutting edge AI has filtered into general applications, often without being called AI because once something becomes useful enough and common enough it's [not labeled AI anymore](#)."^{[2][3]}

Various subfields of AI research are centered around particular goals and the use of particular tools. The traditional goals of AI research include [reasoning](#), [knowledge representation](#), [planning](#), [learning](#), [natural language processing](#), perception, and support for [robotics](#).^[4] [General intelligence](#)—the ability to complete any task performed by a human on an at least equal level—is among the field's long-term goals.^[4] To reach these goals, AI researchers have adapted and integrated a wide range of techniques, including [search](#) and [mathematical optimization](#), [formal logic](#), [artificial neural networks](#), and methods based on [statistics](#), [operations research](#), and [economics](#).^[5] AI also draws upon [psychology](#), [linguistics](#), [philosophy](#), [neuroscience](#), and other fields.^[5]

Artificial intelligence was founded as an academic discipline in 1956,^[6] and the field went through multiple cycles of optimism throughout [its history](#),^{[7][8]} followed by periods of disappointment and loss of funding, known as [AI winters](#).^{[9][10]} Funding and interest vastly increased after 2012 when [deep learning](#) outperformed previous AI techniques.^[11] This growth accelerated further after 2017 with the [transformer architecture](#),^[12] and by the early 2020s many billions of dollars were being invested in AI and the field experienced rapid ongoing [progress](#) in what has become known as the [AI boom](#). The emergence of advanced generative AI in the midst of the AI boom and its ability to create and modify content exposed several unintended consequences and harms in the present and raised concerns about the [risks of AI](#) and [its long-term effects](#) in the future, prompting discussions about [regulatory policies](#) to ensure the [safety and benefits of the technology](#).

Goals

The general problem of simulating (or creating) intelligence has been broken into subproblems. These consist of particular traits or capabilities that researchers expect an intelligent system to display. The traits described below have received the most attention and cover the scope of AI research.^[a]

Reasoning and problem-solving

Early researchers developed algorithms that imitated step-by-step reasoning that humans use when they solve puzzles or make logical [deductions](#).^[13] By the late 1980s and 1990s, methods were developed for dealing with [uncertain](#) or incomplete information, employing concepts from [probability](#) and [economics](#).^[14]

Many of these algorithms are insufficient for solving large reasoning problems because they experience a "combinatorial explosion": They become exponentially slower as the problems grow.^[15] Even humans rarely use the step-by-step deduction that early AI research could model. They solve most of their problems using fast, intuitive judgments.^[16] Accurate and efficient reasoning is an unsolved problem.

Knowledge representation



An ontology represents knowledge as a set of concepts within a domain and the relationships between those concepts.

[Knowledge representation](#) and [knowledge engineering](#)^[17] allow AI programs to answer questions intelligently and make deductions about real-world facts. Formal knowledge representations are used in content-based indexing and retrieval,^[18] scene interpretation,^[19] clinical decision support,^[20] knowledge discovery (mining "interesting" and actionable inferences from large [databases](#)),^[21] and other areas.^[22]

A [knowledge base](#) is a body of knowledge represented in a form that can be used by a program. An [ontology](#) is the set of objects, relations, concepts, and properties used by a particular domain of knowledge.^[23] Knowledge bases need to represent things such as objects, properties, categories, and relations between objects;^[24] situations, events, states, and time;^[25] causes and effects;^[26] knowledge about knowledge (what we know about what other people know);^[27] [default reasoning](#) (things that humans assume are true until they are told differently and will remain true even when other facts are changing);^[28] and many other aspects and domains of knowledge.

Among the most difficult problems in knowledge representation are the breadth of commonsense knowledge (the set of atomic facts that the average person knows is enormous);^[29] and the sub-symbolic form of most commonsense knowledge (much of what people know is not represented as "facts" or "statements" that they could express verbally).^[16] There is also the difficulty of [knowledge acquisition](#), the problem of obtaining knowledge for AI applications.^[c]

Planning and decision-making

An "agent" is anything that perceives and takes actions in the world. A [rational agent](#) has goals or preferences and takes actions to make them happen.^{[d][32]} In [automated planning](#), the agent has a specific goal.^[33] In [automated decision-making](#), the agent has preferences—there are some situations it would prefer to be in, and some situations it is trying to avoid. The decision-making agent assigns a number to each situation (called the "[utility](#)") that measures how much the agent prefers it. For each possible action, it can calculate the "[expected utility](#)": the [utility](#) of all possible outcomes of the action, weighted by the probability that the outcome will occur. It can then choose the action with the maximum expected utility.^[34]

In [classical planning](#), the agent knows exactly what the effect of any action will be.^[35] In most real-world problems, however, the agent may not be certain about the situation they are in (it is "unknown" or "unobservable") and it may not know for certain what will happen after each possible

action (it is not "deterministic"). It must choose an action by making a probabilistic guess and then reassess the situation to see if the action worked.^[36]

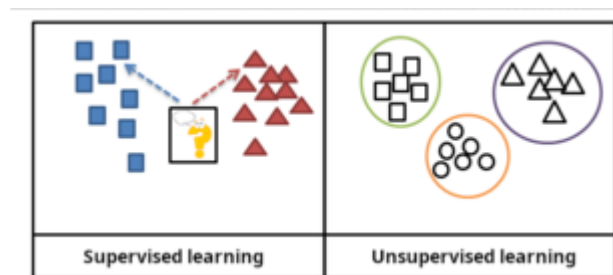
In some problems, the agent's preferences may be uncertain, especially if there are other agents or humans involved. These can be learned (e.g., with [inverse reinforcement learning](#)), or the agent can seek information to improve its preferences.^[37] [Information value theory](#) can be used to weigh the value of exploratory or experimental actions.^[38] The space of possible future actions and situations is typically [intractably](#) large, so the agents must take actions and evaluate situations while being uncertain of what the outcome will be.

A [Markov decision process](#) has a [transition model](#) that describes the probability that a particular action will change the state in a particular way and a [reward function](#) that supplies the utility of each state and the cost of each action. A [policy](#) associates a decision with each possible state. The policy could be calculated (e.g., by [iteration](#)), be [heuristic](#), or it can be learned.^[39]

[Game theory](#) describes the rational behavior of multiple interacting agents and is used in AI programs that make decisions that involve other agents.^[40]

Learning

[Machine learning](#) is the study of programs that can improve their performance on a given task automatically.^[41] It has been a part of AI from the beginning.^[6]



There are several kinds of machine learning. [Unsupervised learning](#) analyzes a stream of data and finds patterns and makes predictions without any other guidance.^[44] [Supervised learning](#) requires labeling the training data with the expected answers, and comes in two main varieties: [classification](#) (where the program must learn to predict what category the input belongs in) and [regression](#) (where the program must deduce a numeric function based on numeric input).^[45]

In [reinforcement learning](#), the agent is rewarded for good responses and punished for bad ones. The agent learns to choose responses that are classified as "good".^[46] [Transfer learning](#) is when the knowledge gained from one problem is applied to a new problem.^[47] [Deep learning](#) is a type of machine learning that runs inputs through biologically inspired [artificial neural networks](#) for all of these types of learning.^[48]

[Computational learning theory](#) can assess learners by [computational complexity](#), by [sample complexity](#) (how much data is required), or by other notions of [optimization](#).^[49]

Natural language processing

[Natural language processing](#) (NLP)^[50] allows programs to read, write and communicate in human languages such as [English](#). Specific problems include [speech recognition](#), [speech synthesis](#), [machine translation](#), [information extraction](#), [information retrieval](#) and [question answering](#).^[51]

Early work, based on [Noam Chomsky's generative grammar](#) and [semantic networks](#), had difficulty with [word-sense disambiguation](#)^[f] unless restricted to small domains called "[micro-worlds](#)" (due to the common sense knowledge problem^[29]). [Margaret Masterman](#) believed that it was meaning and not grammar that was the key to understanding languages, and that [thesauri](#) and not dictionaries should be the basis of computational language structure.

Modern deep learning techniques for NLP include [word embedding](#) (representing words, typically as [vectors](#) encoding their meaning),^[52] [transformers](#) (a deep learning architecture using an [attention](#) mechanism),^[53] and others.^[54] In 2019, [generative pre-trained transformer](#) (or "GPT") language models began to generate coherent text,^{[55][56]} and by 2023, these models were able to get human-level scores on the [bar exam](#), [SAT](#) test, [GRE](#) test, and many other real-world applications.^[57]

Perception

[Machine perception](#) is the ability to use input from sensors (such as cameras, microphones, wireless signals, active [lidar](#), sonar, radar, and [tactile sensors](#)) to deduce aspects of the world. [Computer vision](#) is the ability to analyze visual input.^[58]

The field includes [speech recognition](#),^[59] [image classification](#),^[60] [facial recognition](#), [object recognition](#),^[61] [object tracking](#),^[62] and [robotic perception](#).^[63]

Social intelligence



[Kismet](#), a robot head which was made in the 1990s; it is a machine that can recognize and simulate emotions.^[64]

[Affective computing](#) is a field that comprises systems that recognize, interpret, process, or simulate human [feeling, emotion, and mood](#).^[65] For example, some [virtual assistants](#) are programmed to speak conversationally or even to banter humorously; it makes them appear more sensitive to the emotional dynamics of human interaction, or to otherwise facilitate [human-computer interaction](#).

However, this tends to give naïve users an unrealistic conception of the intelligence of existing computer agents.^[66] Moderate successes related to affective computing include textual [sentiment analysis](#) and, more recently, [multimodal sentiment analysis](#), wherein AI classifies the effects displayed by a videotaped subject.^[67]

General intelligence

A machine with [artificial general intelligence](#) should be able to solve a wide variety of problems with breadth and versatility similar to [human intelligence](#).^[4]

Techniques

AI research uses a wide variety of techniques to accomplish the goals above.^[b]

Search and optimization

AI can solve many problems by intelligently searching through many possible solutions.^[68] There are two very different kinds of search used in AI: [state space search](#) and [local search](#).

State space search

[State space search](#) searches through a tree of possible states to try to find a goal state.^[69] For example, [planning](#) algorithms search through trees of goals and subgoals, attempting to find a path to a target goal, a process called [means-ends analysis](#).^[70]

[Simple exhaustive searches](#)^[71] are rarely sufficient for most real-world problems: the [search space](#) (the number of places to search) quickly grows to [astronomical numbers](#). The result is a search that is [too slow](#) or never completes.^[15] "[Heuristics](#)" or "rules of thumb" can help prioritize choices that are more likely to reach a goal.^[72]

[Adversarial search](#) is used for [game-playing](#) programs, such as chess or Go. It searches through a [tree](#) of possible moves and countermoves, looking for a winning position.^[73]

Local search

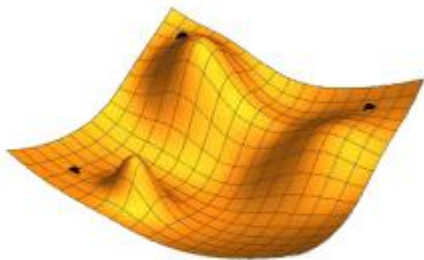


Illustration of [gradient descent](#) for 3 different starting points; two parameters (represented by the plan coordinates) are adjusted in order to minimize the [loss function](#) (the height)

[Local search](#) uses [mathematical optimization](#) to find a solution to a problem. It begins with some form of guess and refines it incrementally.^[74]

[Gradient descent](#) is a type of local search that optimizes a set of numerical parameters by incrementally adjusting them to minimize a [loss function](#). Variants of gradient descent are commonly used to train [neural networks](#),^[75] through the [backpropagation](#) algorithm.

Another type of local search is [evolutionary computation](#), which aims to iteratively improve a set of candidate solutions by "mutating" and "recombining" them, [selecting](#) only the fittest to survive each generation.^[76]

Distributed search processes can coordinate via [swarm intelligence](#) algorithms. Two popular swarm algorithms used in search are [particle swarm optimization](#) (inspired by bird [flocking](#)) and [ant colony optimization](#) (inspired by [ant trails](#)).^[77]

Logic

Formal [logic](#) is used for [reasoning](#) and [knowledge representation](#).^[78] Formal logic comes in two main forms: [propositional logic](#) (which operates on statements that are true or false and uses [logical connectives](#) such as "and", "or", "not" and "implies")^[79] and [predicate logic](#) (which also operates on objects, predicates and relations and uses [quantifiers](#) such as "Every X is a Y" and "There are *some* Xs that are Ys").^[80]

[Deductive reasoning](#) in logic is the process of [proving](#) a new statement ([conclusion](#)) from other statements that are given and assumed to be true (the [premises](#)).^[81] Proofs can be structured as proof [trees](#), in which nodes are labelled by sentences, and children nodes are connected to parent nodes by [inference rules](#).

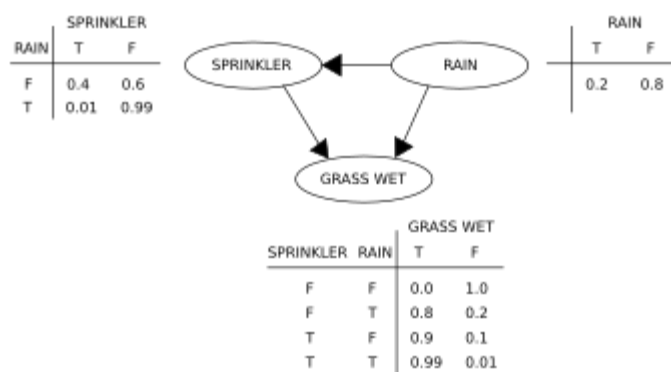
Given a problem and a set of premises, problem-solving reduces to searching for a proof tree whose root node is labelled by a solution of the problem and whose [leaf nodes](#) are labelled by premises or [axioms](#). In the case of [Horn clauses](#), problem-solving search can be performed by reasoning [forwards](#) from the premises or [backwards](#) from the problem.^[82] In the more general case of the clausal form of [first-order logic](#), [resolution](#) is a single, axiom-free rule of inference, in which a problem is solved by proving a contradiction from premises that include the negation of the problem to be solved.^[83]

Inference in both Horn clause logic and first-order logic is [undecidable](#), and therefore [intractable](#). However, backward reasoning with Horn clauses, which underpins computation in the [logic programming](#) language [Prolog](#), is [Turing complete](#). Moreover, its efficiency is competitive with computation in other [symbolic programming](#) languages.^[84]

[Fuzzy logic](#) assigns a "degree of truth" between 0 and 1. It can therefore handle propositions that are vague and partially true.^[85]

[Non-monotonic logics](#), including logic programming with [negation as failure](#), are designed to handle [default reasoning](#).^[28] Other specialized versions of logic have been developed to describe many complex domains.

Probabilistic methods for uncertain reasoning



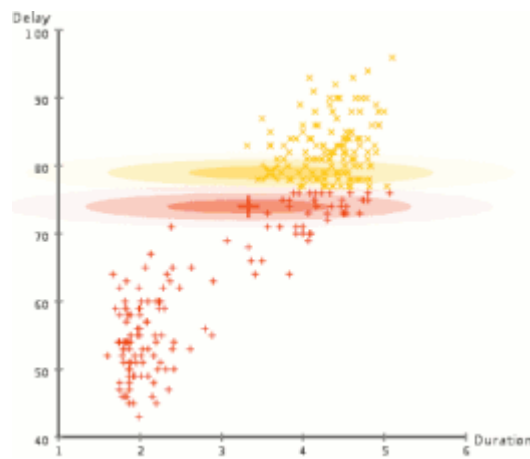
A simple [Bayesian network](#), with the associated [conditional probability tables](#)

Many problems in AI (including in reasoning, planning, learning, perception, and robotics) require the agent to operate with incomplete or uncertain information. AI researchers have devised a number of tools to solve these problems using methods from [probability](#) theory and economics.^[86] Precise mathematical tools have been developed that analyze how an agent can make choices and plan, using [decision theory](#), [decision analysis](#),^[87] and [information value theory](#).^[88] These tools include models such as [Markov decision processes](#),^[89] dynamic [decision networks](#),^[90] [game theory](#) and [mechanism design](#).^[91]

[Bayesian networks](#)^[92] are a tool that can be used for [reasoning](#) (using the [Bayesian inference](#) algorithm),^{[8][94]} [learning](#) (using the [expectation-maximization](#)

[algorithm](#)),^[96] [planning](#) (using [decision networks](#))^[97] and [perception](#) (using [dynamic Bayesian networks](#)).^[90]

Probabilistic algorithms can also be used for filtering, prediction, smoothing, and finding explanations for streams of data, thus helping perception systems analyze processes that occur over time (e.g., [hidden Markov models](#) or [Kalman filters](#)).^[90]



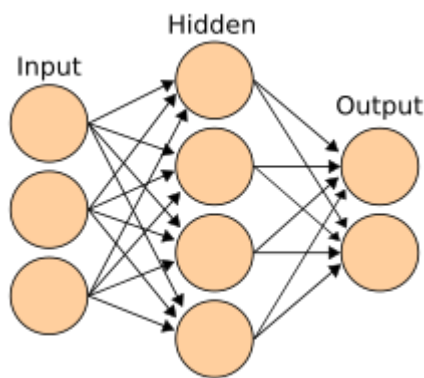
[Expectation–maximization clustering](#) of [Old Faithful](#) eruption data starts from a random guess but then successfully converges on an accurate clustering of the two physically distinct modes of eruption.

Classifiers and statistical learning methods

The simplest AI applications can be divided into two types: classifiers (e.g., "if shiny then diamond"), on one hand, and controllers (e.g., "if diamond then pick up"), on the other hand. [Classifiers](#)^[98] are functions that use [pattern matching](#) to determine the closest match. They can be fine-tuned based on chosen examples using [supervised learning](#). Each pattern (also called an "[observation](#)") is labeled with a certain predefined class. All the observations combined with their class labels are known as a [data set](#). When a new observation is received, that observation is classified based on previous experience.^[45]

There are many kinds of classifiers in use.^[99] The [decision tree](#) is the simplest and most widely used symbolic machine learning algorithm.^[100] [K-nearest neighbor](#) algorithm was the most widely used analogical AI until the mid-1990s, and [kernel methods](#) such as the [support vector machine](#) (SVM) displaced k-nearest neighbor in the 1990s.^[101] The [naive Bayes classifier](#) is reportedly the "most widely used learner"^[102] at Google, due in part to its scalability.^[103] [Neural networks](#) are also used as classifiers.^[104]

Artificial neural networks



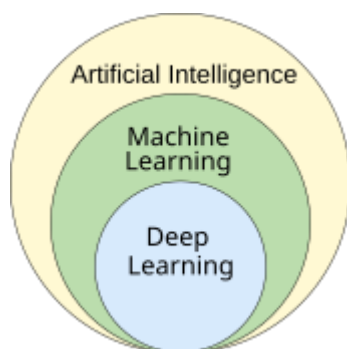
A neural network is an interconnected group of nodes, akin to the vast network of [neurons](#) in the [human brain](#).

An artificial neural network is based on a collection of nodes also known as [artificial neurons](#), which loosely model the [neurons](#) in a biological brain. It is trained to recognise patterns; once trained, it can recognise those patterns in fresh data. There is an input, at least one hidden layer of nodes and an output. Each node applies a function and once the [weight](#) crosses its specified threshold, the data is transmitted to the next layer. A network is typically called a deep neural network if it has at least 2 hidden layers.^[104]

Learning algorithms for neural networks use [local search](#) to choose the weights that will get the right output for each input during training. The most common training technique is the [backpropagation](#) algorithm.^[105] Neural networks learn to model complex relationships between inputs and outputs and [find patterns](#) in data. In theory, a neural network can learn any function.^[106]

In [feedforward neural networks](#) the signal passes in only one direction.^[107] [Recurrent neural networks](#) feed the output signal back into the input, which allows short-term memories of previous input events. [Long short term memory](#) is the most successful network architecture for recurrent networks.^[108] [Perceptrons](#)^[109] use only a single layer of neurons; deep learning^[110] uses multiple layers. [Convolutional neural networks](#) strengthen the connection between neurons that are "close" to each other—this is especially important in [image processing](#), where a local set of neurons must [identify an "edge"](#) before the network can identify an object.^[111]

Deep learning



[Deep learning](#)^[110] uses several layers of neurons between the network's inputs and outputs. The multiple layers can progressively extract higher-level features from the raw input. For example, in [image processing](#), lower layers may identify edges, while higher layers may identify the concepts relevant to a human such as digits, letters, or faces.^[112]

Deep learning has profoundly improved the performance of programs in many important subfields of artificial intelligence, including [computer vision](#), [speech recognition](#), [natural language](#)

[processing](#), [image classification](#),^[113] and others. The reason that deep learning performs so well in so many applications is not known as of 2021.^[114] The sudden success of deep learning in 2012–2015 did not occur because of some new discovery or theoretical breakthrough (deep neural networks and backpropagation had been described by many people, as far back as the 1950s)^[11] but because of two factors: the incredible increase in computer power (including the hundred-fold increase in speed by switching to [GPUs](#)) and the availability of vast amounts of training data, especially the giant [curated datasets](#) used for benchmark testing, such as [ImageNet](#).^[11]

GPT

[Generative pre-trained transformers](#) (GPT) are [large language models](#) (LLMs) that generate text based on the semantic relationships between words in sentences. Text-based GPT models are pretrained on a large [corpus of text](#) that can be from the Internet. The pretraining consists of predicting the next [token](#) (a token being usually a word, subword, or punctuation). Throughout this pretraining, GPT models accumulate knowledge about the world and can then generate human-like text by repeatedly predicting the next token. Typically, a subsequent training phase makes the model more truthful, useful, and harmless, usually with a technique called [reinforcement learning from human feedback](#) (RLHF). Current GPT models are prone to generating falsehoods called "[hallucinations](#)", although this can be reduced with RLHF and quality data. They are used in [chatbots](#), which allow people to ask a question or request a task in simple text.^{[122][123]}

Current models and services include [Gemini](#) (formerly Bard), [ChatGPT](#), [Grok](#), [Claude](#), [Copilot](#), and [LLaMA](#).^[124] [Multimodal](#) GPT models can process different types of data ([modalities](#)) such as images, videos, sound, and text.^[125]

Hardware and software

Main articles: [Programming languages for artificial intelligence](#) and [Hardware for artificial intelligence](#)

In the late 2010s, [graphics processing units](#) (GPUs) that were increasingly designed with AI-specific enhancements and used with specialized [TensorFlow](#) software had replaced previously used [central processing unit](#) (CPUs) as the dominant means for large-scale (commercial and academic) [machine learning](#) models' training.^[126] Specialized [programming languages](#) such as [Prolog](#) were used in early AI research,^[127] but [general-purpose programming languages](#) like [Python](#) have become predominant.^[128]

The transistor density in [integrated circuits](#) has been observed to roughly double every 18 months—a trend known as [Moore's law](#), named after the [Intel](#) co-founder [Gordon Moore](#), who first identified it. Improvements in [GPUs](#) have been even faster,^[129] a trend sometimes called [Huang's law](#),^[130] named after [Nvidia](#) co-founder and CEO [Jensen Huang](#).

Applications

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

AI and machine learning technology is used in most of the essential applications of the 2020s, including: [search engines](#) (such as [Google Search](#)), [targeting online advertisements](#), [recommendation systems](#) (offered by [Netflix](#), [YouTube](#) or [Amazon](#)), driving [internet traffic](#), [targeted advertising](#) ([AdSense](#), [Facebook](#)), [virtual assistants](#) (such as [Siri](#) or [Alexa](#)), [autonomous vehicles](#) (including [drones](#), [ADAS](#) and [self-driving cars](#)), [automatic language translation](#) ([Microsoft Translator](#), [Google Translate](#)), [facial recognition](#) ([Apple's Face](#)

[ID](#) or [Microsoft's DeepFace](#) and [Google's FaceNet](#)) and [image labeling](#) (used by [Facebook](#), Apple's [iPhoto](#) and [TikTok](#)). The deployment of AI may be overseen by a [Chief automation officer](#) (CAO).

Health and medicine

Main article: [Artificial intelligence in healthcare](#)

The application of AI in [medicine](#) and [medical research](#) has the potential to increase patient care and quality of life.^[131] Through the lens of the [Hippocratic Oath](#), medical professionals are ethically compelled to use AI, if applications can more accurately diagnose and treat patients.^{[132][133]}

For medical research, AI is an important tool for processing and integrating [big data](#). This is particularly important for [organoid](#) and [tissue engineering](#) development which use [microscopy](#) imaging as a key technique in fabrication.^[134] It has been suggested that AI can overcome discrepancies in funding allocated to different fields of research.^[134] New AI tools can deepen the understanding of biomedically relevant pathways. For example, [AlphaFold 2](#) (2021) demonstrated the ability to approximate, in hours rather than months, the 3D [structure of a protein](#).^[135] In 2023, it was reported that AI-guided drug discovery helped find a class of antibiotics capable of killing two different types of drug-resistant bacteria.^[136] In 2024, researchers used machine learning to accelerate the search for [Parkinson's disease](#) drug treatments. Their aim was to identify compounds that block the clumping, or aggregation, of [alpha-synuclein](#) (the protein that characterises Parkinson's disease). They were able to speed up the initial screening process ten-fold and reduce the cost by a thousand-fold.^{[137][138]}

Games

Main article: [Game artificial intelligence](#)

[Game playing](#) programs have been used since the 1950s to demonstrate and test AI's most advanced techniques.^[139] [Deep Blue](#) became the first computer chess-playing system to beat a reigning world chess champion, [Garry Kasparov](#), on 11 May 1997.^[140] In 2011, in a [Jeopardy! quiz show](#) exhibition match, [IBM's question answering system, Watson](#), defeated the two greatest *Jeopardy!* champions, [Brad Rutter](#) and [Ken Jennings](#), by a significant margin.^[141] In March 2016, [AlphaGo](#) won 4 out of 5 games of [Go](#) in a match with Go champion [Lee Sedol](#), becoming the first [computer Go](#)-playing system to beat a professional Go player without [handicaps](#). Then, in 2017, it [defeated Ke Jie](#), who was the best Go player in the world.^[142] Other programs handle [imperfect-information](#) games, such as the [poker](#)-playing program [Pluribus](#).^[143] [DeepMind](#) developed increasingly generalistic [reinforcement learning](#) models, such as with [MuZero](#), which could be trained to play chess, Go, or [Atari](#) games.^[144] In 2019, DeepMind's AlphaStar achieved grandmaster level in [StarCraft II](#), a particularly challenging real-time strategy game that involves incomplete knowledge of what happens on the map.^[145] In 2021, an AI agent competed in a PlayStation [Gran Turismo](#) competition, winning against four of the world's best Gran Turismo drivers using deep reinforcement learning.^[146] In 2024, Google DeepMind introduced SIMA, a type of AI capable of autonomously playing nine previously unseen [open-world](#) video games by observing screen output, as well as executing short, specific tasks in response to natural language instructions.^[147]

Mathematics

Large language models, such as [GPT-4](#), [Gemini](#), [Claude](#), [LLaMa](#) or [Mistral](#), are increasingly used in mathematics. These probabilistic models are versatile, but can also produce wrong answers in the form of [hallucinations](#). They sometimes need a large database of mathematical problems to learn

from, but also methods such as [supervised fine-tuning](#)^[148] or trained [classifiers](#) with human-annotated data to improve answers for new problems and learn from corrections.^[149] A February 2024 study showed that the performance of some language models for reasoning capabilities in solving math problems not included in their training data was low, even for problems with only minor deviations from trained data.^[150] One technique to improve their performance involves training the models to produce correct [reasoning](#) steps, rather than just the correct result.^[151] The [Alibaba Group](#) developed a version of its [Qwen](#) models called *Qwen2-Math*, that achieved state-of-the-art performance on several mathematical benchmarks, including 84% accuracy on the MATH dataset of competition mathematics problems.^[152] In January 2025, Microsoft proposed the technique *rStar-Math* that leverages [Monte Carlo tree search](#) and step-by-step reasoning, enabling a relatively small language model like *Qwen-7B* to solve 53% of the [AIME](#) 2024 and 90% of the MATH benchmark problems.^[153]

Alternatively, dedicated models for mathematical problem solving with higher precision for the outcome including proof of theorems have been developed such as *AlphaTensor*, [AlphaGeometry](#) and *AlphaProof* all from [Google DeepMind](#),^[154] *Llemma* from [EleutherAI](#)^[155] or *Julius*.^[156]

When natural language is used to describe mathematical problems, converters can transform such prompts into a formal language such as [Lean](#) to define mathematical tasks.

Some models have been developed to solve challenging problems and reach good results in benchmark tests, others to serve as educational tools in mathematics.^[157]

[Topological deep learning](#) integrates various [topological](#) approaches.

Finance

Finance is one of the fastest growing sectors where applied AI tools are being deployed: from retail online banking to investment advice and insurance, where automated "robot advisers" have been in use for some years.^[158]

[World Pensions](#) experts like Nicolas Firzli insist it may be too early to see the emergence of highly innovative AI-informed financial products and services: "the deployment of AI tools will simply further automatise things: destroying tens of thousands of jobs in banking, financial planning, and pension advice in the process, but I'm not sure it will unleash a new wave of [e.g., sophisticated] pension innovation."^[159]

Military

Main article: [Military applications of artificial intelligence](#)

Various countries are deploying AI military applications.^[160] The main applications enhance [command and control](#), communications, sensors, integration and interoperability.^[161] Research is targeting intelligence collection and analysis, logistics, cyber operations, information operations, and semiautonomous and [autonomous vehicles](#).^[160] AI technologies enable coordination of sensors and effectors, threat detection and identification, marking of enemy positions, [target acquisition](#), coordination and deconfliction of distributed [Joint Fires](#) between networked combat vehicles involving manned and unmanned teams.^[161]

AI has been used in military operations in Iraq, Syria, Israel and Ukraine.^{[160][162][163][164]}

Generative AI



[Vincent van Gogh](#) in watercolour created by generative AI

software

These paragraphs are an excerpt from [Generative artificial intelligence](#).[\[edit\]](#)

[Generative artificial intelligence](#) (generative AI, GenAI,^{[\[165\]](#)} or GAI) is a subset of artificial intelligence that uses generative models to produce text, images, videos, or other forms of data.^{[\[166\]](#)[\[167\]](#)[\[168\]](#)} These models [learn](#) the underlying patterns and structures of their [training data](#) and use them to produce new data^{[\[169\]](#)[\[170\]](#)} based on the input, which often comes in the form of natural language [prompts](#).^{[\[171\]](#)[\[172\]](#)}

Improvements in [transformer](#)-based [deep neural networks](#), particularly [large language models](#) (LLMs), enabled an [AI boom](#) of generative AI systems in the 2020s. These include [chatbots](#) such as [ChatGPT](#), [Copilot](#), [Gemini](#), and [LLaMA](#); [text-to-image artificial intelligence](#) [image generation](#) systems such as [Stable Diffusion](#), [Midjourney](#), and [DALL-E](#); and [text-to-video](#) AI generators such as [Sora](#).^{[\[173\]](#)[\[174\]](#)[\[175\]](#)[\[176\]](#)} Companies such as [OpenAI](#), [Anthropic](#), [Microsoft](#), [Google](#), and [Baidu](#) as well as numerous smaller firms have developed generative AI models.^{[\[171\]](#)[\[177\]](#)[\[178\]](#)}

Generative AI has uses across a wide range of industries, including software development, healthcare, finance, entertainment, customer service,^{[\[179\]](#)} sales and marketing,^{[\[180\]](#)} art, writing,^{[\[181\]](#)} fashion,^{[\[182\]](#)} and product design.^{[\[183\]](#)} However, concerns have been raised about the potential misuse of generative AI such as [cybercrime](#), the use of [fake news](#) or [deepfakes](#) to deceive or manipulate people, and [the mass replacement of human jobs](#).^{[\[184\]](#)[\[185\]](#)} Intellectual property law concerns also exist around generative models that are trained on and emulate copyrighted works of art.^{[\[186\]](#)}

Agents

Artificial intelligent (AI) agents are software entities designed to perceive their environment, make decisions, and take actions autonomously to achieve specific goals. These agents can interact with users, their environment, or other agents. AI agents are used in various applications, including [virtual assistants](#), [chatbots](#), [autonomous vehicles](#), [game-playing systems](#), and [industrial robotics](#). AI agents operate within the constraints of their programming, available computational resources, and hardware limitations. This means they are restricted to performing tasks within their defined scope and have finite memory and processing capabilities. In real-world applications, AI agents often face time constraints for decision-making and action execution. Many AI agents incorporate learning algorithms, enabling them to improve their performance over time through experience or training. Using machine learning, AI agents can adapt to new situations and optimise their behaviour for their designated tasks.^{[\[187\]](#)[\[188\]](#)[\[189\]](#)}

Sexuality

Applications of AI in this domain include AI-enabled menstruation and fertility trackers that analyze user data to offer prediction,^[190] AI-integrated sex toys (e.g., [teledildonics](#)),^[191] AI-generated sexual education content,^[192] and AI agents that simulate sexual and romantic partners (e.g., [Replika](#)).^[193] AI is also used for the production of non-consensual [deepfake pornography](#), raising significant ethical and legal concerns.^[194]

AI technologies have also been used to attempt to identify [online gender-based violence](#) and online [sexual grooming](#) of minors.^{[195][196]}

Other industry-specific tasks

There are also thousands of successful AI applications used to solve specific problems for specific industries or institutions. In a 2017 survey, one in five companies reported having incorporated "AI" in some offerings or processes.^[197] A few examples are [energy storage](#), medical diagnosis, military logistics, applications that predict the result of judicial decisions, [foreign policy](#), or supply chain management.

AI applications for evacuation and [disaster](#) management are growing. AI has been used to investigate if and how people evacuated in large scale and small scale evacuations using historical data from GPS, videos or social media. Further, AI can provide real time information on the real time evacuation conditions.^{[198][199][200]}

In agriculture, AI has helped farmers identify areas that need irrigation, fertilization, pesticide treatments or increasing yield. Agronomists use AI to conduct research and development. AI has been used to predict the ripening time for crops such as tomatoes, monitor soil moisture, operate agricultural robots, conduct [predictive analytics](#), classify livestock pig call emotions, automate greenhouses, detect diseases and pests, and save water.

Artificial intelligence is used in astronomy to analyze increasing amounts of available data and applications, mainly for "classification, regression, clustering, forecasting, generation, discovery, and the development of new scientific insights." For example, it is used for discovering exoplanets, forecasting solar activity, and distinguishing between signals and instrumental effects in gravitational wave astronomy. Additionally, it could be used for activities in space, such as space exploration, including the analysis of data from space missions, real-time science decisions of spacecraft, space debris avoidance, and more autonomous operation.

During the [2024 Indian elections](#), US\$50 million was spent on authorized AI-generated content, notably by creating [deepfakes](#) of allied (including sometimes deceased) politicians to better engage with voters, and by translating speeches to various local languages.^[201]

Ethics

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

AI has potential benefits and potential risks.^[202] AI may be able to advance science and find solutions for serious problems: [Demis Hassabis](#) of [DeepMind](#) hopes to "solve intelligence, and then use that to solve everything else".^[203] However, as the use of AI has become widespread, several unintended consequences and risks have been identified.^[204] In-production systems can sometimes not factor ethics and bias into their AI training processes, especially when the AI algorithms are inherently unexplainable in deep learning.^[205]

Risks and harm

Privacy and copyright

Further information: [Information privacy](#) and [Artificial intelligence and copyright](#)

Machine learning algorithms require large amounts of data. The techniques used to acquire this data have raised concerns about [privacy](#), [surveillance](#) and [copyright](#).

AI-powered devices and services, such as virtual assistants and IoT products, continuously collect personal information, raising concerns about intrusive data gathering and unauthorized access by third parties. The loss of privacy is further exacerbated by AI's ability to process and combine vast amounts of data, potentially leading to a surveillance society where individual activities are constantly monitored and analyzed without adequate safeguards or transparency.

Sensitive user data collected may include online activity records, geolocation data, video, or audio.^[206] For example, in order to build [speech recognition](#) algorithms, [Amazon](#) has recorded millions of private conversations and allowed [temporary workers](#) to listen to and transcribe some of them.^[207] Opinions about this widespread surveillance range from those who see it as a [necessary evil](#) to those for whom it is clearly [unethical](#) and a violation of the [right to privacy](#).^[208]

AI developers argue that this is the only way to deliver valuable applications and have developed several techniques that attempt to preserve privacy while still obtaining the data, such as [data aggregation](#), [de-identification](#) and [differential privacy](#).^[209] Since 2016, some privacy experts, such as [Cynthia Dwork](#), have begun to view privacy in terms of [fairness](#). [Brian Christian](#) wrote that experts have pivoted "from the question of 'what they know' to the question of 'what they're doing with it'".^[210]

Generative AI is often trained on unlicensed copyrighted works, including in domains such as images or computer code; the output is then used under the rationale of "[fair use](#)". Experts disagree about how well and under what circumstances this rationale will hold up in courts of law; relevant factors may include "the purpose and character of the use of the copyrighted work" and "the effect upon the potential market for the copyrighted work".^{[211][212]} Website owners who do not wish to have their content scraped can indicate it in a "[robots.txt](#)" file.^[213] In 2023, leading authors (including [John Grisham](#) and [Jonathan Franzen](#)) sued AI companies for using their work to train generative AI.^{[214][215]} Another discussed approach is to envision a separate [sui generis](#) system of protection for creations generated by AI to ensure fair attribution and compensation for human authors.^[216]

Dominance by tech giants

The commercial AI scene is dominated by [Big Tech](#) companies such as [Alphabet Inc.](#), [Amazon](#), [Apple Inc.](#), [Meta Platforms](#), and [Microsoft](#).^{[217][218][219]} Some of these players already own the vast majority of existing [cloud infrastructure](#) and [computing](#) power from [data centers](#), allowing them to entrench further in the marketplace.^{[220][221]}

Power needs and environmental impacts

See also: [Environmental impacts of artificial intelligence](#)

In January 2024, the [International Energy Agency](#) (IEA) released *Electricity 2024, Analysis and Forecast to 2026*, forecasting electric power use.^[222] This is the first IEA report to make projections for data centers and power consumption for artificial intelligence and cryptocurrency. The report

states that power demand for these uses might double by 2026, with additional electric power usage equal to electricity used by the whole Japanese nation.^[223]

Prodigious power consumption by AI is responsible for the growth of fossil fuels use, and might delay closings of obsolete, carbon-emitting coal energy facilities. There is a feverish rise in the construction of data centers throughout the US, making large technology firms (e.g., Microsoft, Meta, Google, Amazon) into voracious consumers of electric power. Projected electric consumption is so immense that there is concern that it will be fulfilled no matter the source. A ChatGPT search involves the use of 10 times the electrical energy as a Google search. The large firms are in haste to find power sources – from nuclear energy to geothermal to fusion. The tech firms argue that – in the long view – AI will be eventually kinder to the environment, but they need the energy now. AI makes the power grid more efficient and "intelligent", will assist in the growth of nuclear power, and track overall carbon emissions, according to technology firms.^[224]

A 2024 [Goldman Sachs](#) Research Paper, *AI Data Centers and the Coming US Power Demand Surge*, found "US power demand (is) likely to experience growth not seen in a generation...." and forecasts that, by 2030, US data centers will consume 8% of US power, as opposed to 3% in 2022, presaging growth for the electrical power generation industry by a variety of means.^[225] Data centers' need for more and more electrical power is such that they might max out the electrical grid. The Big Tech companies counter that AI can be used to maximize the utilization of the grid by all.^[226]

In 2024, the *Wall Street Journal* reported that big AI companies have begun negotiations with the US nuclear power providers to provide electricity to the data centers. In March 2024 Amazon purchased a Pennsylvania nuclear-powered data center for \$650 Million (US).^[227] [Nvidia](#) CEO [Jen-Hsun Huang](#) said nuclear power is a good option for the data centers.^[228]

In September 2024, [Microsoft](#) announced an agreement with [Constellation Energy](#) to re-open the [Three Mile Island](#) nuclear power plant to provide Microsoft with 100% of all electric power produced by the plant for 20 years. Reopening the plant, which suffered a partial nuclear meltdown of its Unit 2 reactor in 1979, will require Constellation to get through strict regulatory processes which will include extensive safety scrutiny from the US [Nuclear Regulatory Commission](#). If approved (this will be the first ever US re-commissioning of a nuclear plant), over 835 megawatts of power – enough for 800,000 homes – of energy will be produced. The cost for re-opening and upgrading is estimated at \$1.6 billion (US) and is dependent on tax breaks for nuclear power contained in the 2022 US [Inflation Reduction Act](#).^[229] The US government and the state of Michigan are investing almost \$2 billion (US) to reopen the [Palisades Nuclear](#) reactor on Lake Michigan. Closed since 2022, the plant is planned to be reopened in October 2025. The Three Mile Island facility will be renamed the Crane Clean Energy Center after Chris Crane, a nuclear proponent and former CEO of [Exelon](#) who was responsible for Exelon spinoff of Constellation.^[230]

After the last approval in September 2023, [Taiwan](#) suspended the approval of data centers north of [Taoyuan](#) with a capacity of more than 5 MW in 2024, due to power supply shortages.^[231] Taiwan aims to [phase out nuclear power](#) by 2025.^[231] On the other hand, [Singapore](#) imposed a ban on the opening of data centers in 2019 due to electric power, but in 2022, lifted this ban.^[231]

Although most nuclear plants in Japan have been shut down after the 2011 [Fukushima nuclear accident](#), according to an October 2024 *Bloomberg* article in Japanese, cloud gaming services company Ubitus, in which Nvidia has a stake, is looking for land in Japan near nuclear power plant for a new data center for generative AI.^[232] Ubitus CEO Wesley Kuo said nuclear power plants are the most efficient, cheap and stable power for AI.^[232]

On 1 November 2024, the [Federal Energy Regulatory Commission](#) (FERC) rejected an application submitted by [Talen Energy](#) for approval to supply some electricity from the nuclear power station [Susquehanna](#) to Amazon's data center.^[233] According to the Commission Chairman [Willie L. Phillips](#), it is a burden on the electricity grid as well as a significant cost shifting concern to households and other business sectors.^[233]

Misinformation

See also: [YouTube & Moderation and offensive content](#)

[YouTube](#), [Facebook](#) and others use [recommender systems](#) to guide users to more content. These AI programs were given the goal of [maximizing](#) user engagement (that is, the only goal was to keep people watching). The AI learned that users tended to choose [misinformation](#), [conspiracy theories](#), and extreme [partisan](#) content, and, to keep them watching, the AI recommended more of it. Users also tended to watch more content on the same subject, so the AI led people into [filter bubbles](#) where they received multiple versions of the same misinformation.^[234] This convinced many users that the misinformation was true, and ultimately undermined trust in institutions, the media and the government.^[235] The AI program had correctly learned to maximize its goal, but the result was harmful to society. After the U.S. election in 2016, major technology companies took steps to mitigate the problem ^[citation needed].

In 2022, [generative AI](#) began to create images, audio, video and text that are indistinguishable from real photographs, recordings, films, or human writing. It is possible for bad actors to use this technology to create massive amounts of misinformation or propaganda.^[236] AI pioneer [Geoffrey Hinton](#) expressed concern about AI enabling "authoritarian leaders to manipulate their electorates" on a large scale, among other risks.^[237]

Algorithmic bias and fairness

Main articles: [Algorithmic bias](#) and [Fairness \(machine learning\)](#)

Machine learning applications will be [biased](#)^[k] if they learn from biased data.^[239] The developers may not be aware that the bias exists.^[240] Bias can be introduced by the way [training data](#) is selected and by the way a model is deployed.^{[241][239]} If a biased algorithm is used to make decisions that can seriously [harm](#) people (as it can in [medicine](#), [finance](#), [recruitment](#), [housing](#) or [policing](#)) then the algorithm may cause [discrimination](#).^[242] The field of [fairness](#) studies how to prevent harms from algorithmic biases.

On June 28, 2015, [Google Photos](#)'s new image labeling feature mistakenly identified Jacky Alcine and a friend as "gorillas" because they were black. The system was trained on a dataset that contained very few images of black people,^[243] a problem called "sample size disparity".^[244] Google "fixed" this problem by preventing the system from labelling *anything* as a "gorilla". Eight years later, in 2023, Google Photos still could not identify a gorilla, and neither could similar products from Apple, Facebook, Microsoft and Amazon.^[245]

[COMPAS](#) is a commercial program widely used by [U.S. courts](#) to assess the likelihood of a [defendant](#) becoming a [recidivist](#). In 2016, [Julia Angwin](#) at [ProPublica](#) discovered that COMPAS exhibited racial bias, despite the fact that the program was not told the races of the defendants. Although the error rate for both whites and blacks was calibrated equal at exactly 61%, the errors for each race were different—the system consistently overestimated the chance that a black person would re-offend and would underestimate the chance that a white person would not re-offend.^[246] In 2017, several researchers^[l] showed that it was mathematically impossible for COMPAS to

accommodate all possible measures of fairness when the base rates of re-offense were different for whites and blacks in the data.^[248]

A program can make biased decisions even if the data does not explicitly mention a problematic feature (such as "race" or "gender"). The feature will correlate with other features (like "address", "shopping history" or "first name"), and the program will make the same decisions based on these features as it would on "race" or "gender".^[249] Moritz Hardt said "the most robust fact in this research area is that fairness through blindness doesn't work."^[250]

Criticism of COMPAS highlighted that machine learning models are designed to make "predictions" that are only valid if we assume that the future will resemble the past. If they are trained on data that includes the results of racist decisions in the past, machine learning models must predict that racist decisions will be made in the future. If an application then uses these predictions as *recommendations*, some of these "recommendations" will likely be racist.^[251] Thus, machine learning is not well suited to help make decisions in areas where there is hope that the future will be *better* than the past. It is descriptive rather than prescriptive.^[m]

Bias and unfairness may go undetected because the developers are overwhelmingly white and male: among AI engineers, about 4% are black and 20% are women.^[244]

There are various conflicting definitions and mathematical models of fairness. These notions depend on ethical assumptions, and are influenced by beliefs about society. One broad category is [distributive fairness](#), which focuses on the outcomes, often identifying groups and seeking to compensate for statistical disparities. Representational fairness tries to ensure that AI systems do not reinforce negative [stereotypes](#) or render certain groups invisible. Procedural fairness focuses on the decision process rather than the outcome. The most relevant notions of fairness may depend on the context, notably the type of AI application and the stakeholders. The subjectivity in the notions of bias and fairness makes it difficult for companies to operationalize them. Having access to sensitive attributes such as race or gender is also considered by many AI ethicists to be necessary in order to compensate for biases, but it may conflict with [anti-discrimination laws](#).^[238]

At its 2022 [Conference on Fairness, Accountability, and Transparency](#) (ACM FAccT 2022), the [Association for Computing Machinery](#), in Seoul, South Korea, presented and published findings that recommend that until AI and robotics systems are demonstrated to be free of bias mistakes, they are unsafe, and the use of self-learning neural networks trained on vast, unregulated sources of flawed internet data should be curtailed.^{[dubious – discuss][253]}

Lack of transparency

See also: [Explainable AI](#), [Algorithmic transparency](#), and [Right to explanation](#)

Many AI systems are so complex that their designers cannot explain how they reach their decisions.^[254] Particularly with [deep neural networks](#), in which there are a large amount of non-[linear](#) relationships between inputs and outputs. But some popular explainability techniques exist.^[255]

It is impossible to be certain that a program is operating correctly if no one knows how exactly it works. There have been many cases where a machine learning program passed rigorous tests, but nevertheless learned something different than what the programmers intended. For example, a system that could identify skin diseases better than medical professionals was found to actually have a strong tendency to classify images with a [ruler](#) as "cancerous", because pictures of malignancies typically include a ruler to show the scale.^[256] Another machine learning system designed to help

effectively allocate medical resources was found to classify patients with asthma as being at "low risk" of dying from pneumonia. Having asthma is actually a severe risk factor, but since the patients having asthma would usually get much more medical care, they were relatively unlikely to die according to the training data. The correlation between asthma and low risk of dying from pneumonia was real, but misleading.^[257]

People who have been harmed by an algorithm's decision have a right to an explanation.^[258] Doctors, for example, are expected to clearly and completely explain to their colleagues the reasoning behind any decision they make. Early drafts of the European Union's [General Data Protection Regulation](#) in 2016 included an explicit statement that this right exists.^[6] Industry experts noted that this is an unsolved problem with no solution in sight. Regulators argued that nevertheless the harm is real: if the problem has no solution, the tools should not be used.^[259]

[DARPA](#) established the [XAI](#) ("Explainable Artificial Intelligence") program in 2014 to try to solve these problems.^[260]

Several approaches aim to address the transparency problem. SHAP enables to visualise the contribution of each feature to the output.^[261] LIME can locally approximate a model's outputs with a simpler, interpretable model.^[262] [Multitask learning](#) provides a large number of outputs in addition to the target classification. These other outputs can help developers deduce what the network has learned.^[263] [Deconvolution](#), [DeepDream](#) and other [generative](#) methods can allow developers to see what different layers of a deep network for computer vision have learned, and produce output that can suggest what the network is learning.^[264] For [generative pre-trained transformers](#), [Anthropic](#) developed a technique based on [dictionary learning](#) that associates patterns of neuron activations with human-understandable concepts.^[265]

Bad actors and weaponized AI

Main articles: [Lethal autonomous weapon](#), [Artificial intelligence arms race](#), and [AI safety](#)

Artificial intelligence provides a number of tools that are useful to [bad actors](#), such as [authoritarian governments](#), [terrorists](#), [criminals](#) or [rogue states](#).

A lethal autonomous weapon is a machine that locates, selects and engages human targets without human supervision.^[6] Widely available AI tools can be used by bad actors to develop inexpensive autonomous weapons and, if produced at scale, they are potentially [weapons of mass destruction](#).^[267] Even when used in conventional warfare, they currently cannot reliably choose targets and could potentially [kill an innocent person](#).^[267] In 2014, 30 nations (including China) supported a ban on autonomous weapons under the [United Nations' Convention on Certain Conventional Weapons](#), however the [United States](#) and others disagreed.^[268] By 2015, over fifty countries were reported to be researching battlefield robots.^[269]

AI tools make it easier for [authoritarian governments](#) to efficiently control their citizens in several ways. [Face](#) and [voice recognition](#) allow widespread [surveillance](#). [Machine learning](#), operating this data, can [classify](#) potential enemies of the state and prevent them from hiding. [Recommendation systems](#) can precisely target [propaganda](#) and [misinformation](#) for maximum effect. [Deepfakes](#) and [generative AI](#) aid in producing misinformation. Advanced AI can make authoritarian [centralized decision making](#) more competitive than liberal and decentralized systems such as [markets](#). It lowers the cost and difficulty of [digital warfare](#) and [advanced spyware](#).^[270] All these technologies have been available since 2020 or earlier—AI [facial recognition systems](#) are already being used for [mass surveillance](#) in China.^{[271][272]}

There many other ways that AI is expected to help bad actors, some of which can not be foreseen. For example, machine-learning AI is able to design tens of thousands of toxic molecules in a matter of hours.^[273]

Technological unemployment

Main articles: [Workplace impact of artificial intelligence](#) and [Technological unemployment](#)

Economists have frequently highlighted the risks of redundancies from AI, and speculated about unemployment if there is no adequate social policy for full employment.^[274]

In the past, technology has tended to increase rather than reduce total employment, but economists acknowledge that "we're in uncharted territory" with AI.^[275] A survey of economists showed disagreement about whether the increasing use of robots and AI will cause a substantial increase in long-term [unemployment](#), but they generally agree that it could be a net benefit if [productivity](#) gains are [redistributed](#).^[276] Risk estimates vary; for example, in the 2010s, Michael Osborne and [Carl Benedikt Frey](#) estimated 47% of U.S. jobs are at "high risk" of potential automation, while an OECD report classified only 9% of U.S. jobs as "high risk".^{[6][278]} The methodology of speculating about future employment levels has been criticised as lacking evidential foundation, and for implying that technology, rather than social policy, creates unemployment, as opposed to redundancies.^[274] In April 2023, it was reported that 70% of the jobs for Chinese video game illustrators had been eliminated by generative artificial intelligence.^{[279][280]}

Unlike previous waves of automation, many middle-class jobs may be eliminated by artificial intelligence; [The Economist](#) stated in 2015 that "the worry that AI could do to white-collar jobs what steam power did to blue-collar ones during the Industrial Revolution" is "worth taking seriously".^[281] Jobs at extreme risk range from [paralegals](#) to fast food cooks, while job demand is likely to increase for care-related professions ranging from personal healthcare to the clergy.^[282]

From the early days of the development of artificial intelligence, there have been arguments, for example, those put forward by [Joseph Weizenbaum](#), about whether tasks that can be done by computers actually should be done by them, given the difference between computers and humans, and between quantitative calculation and qualitative, value-based judgement.^[283]

Existential risk

Main article: [Existential risk from artificial intelligence](#)

It has been argued AI will become so powerful that humanity may irreversibly lose control of it. This could, as physicist [Stephen Hawking](#) stated, "[spell the end of the human race](#)".^[284] This scenario has been common in science fiction, when a computer or robot suddenly develops a human-like "self-awareness" (or "sentience" or "consciousness") and becomes a malevolent character.^[6] These sci-fi scenarios are misleading in several ways.

First, AI does not require human-like [sentience](#) to be an existential risk. Modern AI programs are given specific goals and use learning and intelligence to achieve them. Philosopher [Nick Bostrom](#) argued that if one gives *almost any* goal to a sufficiently powerful AI, it may choose to destroy humanity to achieve it (he used the example of a [paperclip factory manager](#)).^[286] [Stuart Russell](#) gives the example of household robot that tries to find a way to kill its owner to prevent it from being unplugged, reasoning that "you can't fetch the coffee if you're dead."^[287] In order to be safe for humanity, a [superintelligence](#) would have to be genuinely [aligned](#) with humanity's morality and values so that it is "fundamentally on our side".^[288]

Second, [Yuval Noah Harari](#) argues that AI does not require a robot body or physical control to pose an existential risk. The essential parts of civilization are not physical. Things like [ideologies](#), [law](#), [government](#), [money](#) and the [economy](#) are built on [language](#); they exist because there are stories that billions of people believe. The current prevalence of [misinformation](#) suggests that an AI could use language to convince people to believe anything, even to take actions that are destructive.^[289]

The opinions amongst experts and industry insiders are mixed, with sizable fractions both concerned and unconcerned by risk from eventual superintelligent AI.^[290] Personalities such as [Stephen Hawking](#), [Bill Gates](#), and [Elon Musk](#),^[291] as well as AI pioneers such as [Yoshua Bengio](#), [Stuart Russell](#), [Demis Hassabis](#), and [Sam Altman](#), have expressed concerns about existential risk from AI.

In May 2023, [Geoffrey Hinton](#) announced his resignation from Google in order to be able to "freely speak out about the risks of AI" without "considering how this impacts Google".^[292] He notably mentioned risks of an [AI takeover](#),^[293] and stressed that in order to avoid the worst outcomes, establishing safety guidelines will require cooperation among those competing in use of AI.^[294]

In 2023, many leading AI experts endorsed [the joint statement](#) that "Mitigating the risk of extinction from AI should be a global priority alongside other societal-scale risks such as pandemics and nuclear war".^[295]

Some other researchers were more optimistic. AI pioneer [Jürgen Schmidhuber](#) did not sign the joint statement, emphasising that in 95% of all cases, AI research is about making "human lives longer and healthier and easier."^[296] While the tools that are now being used to improve lives can also be used by bad actors, "they can also be used against the bad actors."^{[297][298]} [Andrew Ng](#) also argued that "it's a mistake to fall for the doomsday hype on AI—and that regulators who do will only benefit vested interests."^[299] [Yann LeCun](#) "scoffs at his peers' dystopian scenarios of supercharged misinformation and even, eventually, human extinction."^[300] In the early 2010s, experts argued that the risks are too distant in the future to warrant research or that humans will be valuable from the perspective of a superintelligent machine.^[301] However, after 2016, the study of current and future risks and possible solutions became a serious area of research.^[302]

Ethical machines and alignment

Main articles: [Machine ethics](#), [AI safety](#), [Friendly artificial intelligence](#), [Artificial moral agents](#), and [Human Compatible](#)

Friendly AI are machines that have been designed from the beginning to minimize risks and to make choices that benefit humans. [Eliezer Yudkowsky](#), who coined the term, argues that developing friendly AI should be a higher research priority: it may require a large investment and it must be completed before AI becomes an existential risk.^[303]

Machines with intelligence have the potential to use their intelligence to make ethical decisions. The field of machine ethics provides machines with ethical principles and procedures for resolving ethical dilemmas.^[304] The field of machine ethics is also called computational morality,^[304] and was founded at an [AAAI](#) symposium in 2005.^[305]

Other approaches include [Wendell Wallach](#)'s "artificial moral agents"^[306] and [Stuart J. Russell](#)'s [three principles](#) for developing provably beneficial machines.^[307]

Open source

Active organizations in the AI open-source community include [Hugging Face](#),^[308] [Google](#),^[309] [EleutherAI](#) and [Meta](#).^[310] Various AI models, such as [Llama 2](#), [Mistral](#) or [Stable Diffusion](#), have been made open-weight,^{[311][312]} meaning that their architecture and trained parameters (the "weights") are publicly available. Open-weight models can be freely [fine-tuned](#), which allows companies to specialize them with their own data and for their own use-case.^[313] Open-weight models are useful for research and innovation but can also be misused. Since they can be fine-tuned, any built-in security measure, such as objecting to harmful requests, can be trained away until it becomes ineffective. Some researchers warn that future AI models may develop dangerous capabilities (such as the potential to drastically facilitate [bioterrorism](#)) and that once released on the Internet, they cannot be deleted everywhere if needed. They recommend pre-release audits and cost-benefit analyses.^[314]

Frameworks

Artificial Intelligence projects can have their ethical permissibility tested while designing, developing, and implementing an AI system. An AI framework such as the Care and Act Framework containing the SUM values—developed by the [Alan Turing Institute](#) tests projects in four main areas:^{[315][316]}

- **Respect** the dignity of individual people
- **Connect** with other people sincerely, openly, and inclusively
- **Care** for the wellbeing of everyone
- **Protect** social values, justice, and the public interest

Other developments in ethical frameworks include those decided upon during the [Asilomar Conference](#), the Montreal Declaration for Responsible AI, and the IEEE's Ethics of Autonomous Systems initiative, among others;^[317] however, these principles do not go without their criticisms, especially regards to the people chosen contributes to these frameworks.^[318]

Promotion of the wellbeing of the people and communities that these technologies affect requires consideration of the social and ethical implications at all stages of AI system design, development and implementation, and collaboration between job roles such as data scientists, product managers, data engineers, domain experts, and delivery managers.^[319]

The [UK AI Safety Institute](#) released in 2024 a testing toolset called 'Inspect' for AI safety evaluations available under a MIT open-source licence which is freely available on GitHub and can be improved with third-party packages. It can be used to evaluate AI models in a range of areas including core knowledge, ability to reason, and autonomous capabilities.^[320]

Regulation

Main articles: [Regulation of artificial intelligence](#), [Regulation of algorithms](#), and [AI safety](#)



The first global [AI Safety Summit](#) was held in 2023 with a declaration calling for international cooperation.

The regulation of artificial intelligence is the development of public sector policies and laws for promoting and regulating AI; it is therefore related to the broader regulation of algorithms.^[321] The regulatory and policy landscape for AI is an emerging issue in jurisdictions globally.^[322] According to AI Index at [Stanford](#), the annual number of AI-related laws passed in the 127 survey countries jumped from one passed in 2016 to 37 passed in 2022 alone.^{[323][324]} Between 2016 and 2020, more than 30 countries adopted dedicated strategies for AI.^[325] Most EU member states had released national AI strategies, as had Canada, China, India, Japan, Mauritius, the Russian Federation, Saudi Arabia, United Arab Emirates, U.S., and Vietnam. Others were in the process of elaborating their own AI strategy, including Bangladesh, Malaysia and Tunisia.^[325] The [Global Partnership on Artificial Intelligence](#) was launched in June 2020, stating a need for AI to be developed in accordance with human rights and democratic values, to ensure public confidence and trust in the technology.^[325] [Henry Kissinger](#), [Eric Schmidt](#), and [Daniel Huttenlocher](#) published a joint statement in November 2021 calling for a government commission to regulate AI.^[326] In 2023, OpenAI leaders published recommendations for the governance of superintelligence, which they believe may happen in less than 10 years.^[327] In 2023, the United Nations also launched an advisory body to provide recommendations on AI governance; the body comprises technology company executives, governments officials and academics.^[328] In 2024, the [Council of Europe](#) created the first international legally binding treaty on AI, called the "[Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law](#)". It was adopted by the European Union, the United States, the United Kingdom, and other signatories.^[329]

In a 2022 [Ipsos](#) survey, attitudes towards AI varied greatly by country; 78% of Chinese citizens, but only 35% of Americans, agreed that "products and services using AI have more benefits than drawbacks".^[323] A 2023 [Reuters/Ipsos](#) poll found that 61% of Americans agree, and 22% disagree, that AI poses risks to humanity.^[330] In a 2023 [Fox News](#) poll, 35% of Americans thought it "very important", and an additional 41% thought it "somewhat important", for the federal government to regulate AI, versus 13% responding "not very important" and 8% responding "not at all important".^{[331][332]}

In November 2023, the first global [AI Safety Summit](#) was held in [Bletchley Park](#) in the UK to discuss the near and far term risks of AI and the possibility of mandatory and voluntary regulatory frameworks.^[333] 28 countries including the United States, China, and the European Union issued a declaration at the start of the summit, calling for international co-operation to manage the challenges and risks of artificial intelligence.^{[334][335]} In May 2024 at the [AI Seoul Summit](#), 16 global AI tech companies agreed to safety commitments on the development of AI.^{[336][337]}

History

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

For a chronological guide, see [Timeline of artificial intelligence](#).

The study of mechanical or "formal" reasoning began with philosophers and mathematicians in antiquity. The study of logic led directly to [Alan Turing's theory of computation](#), which suggested that a machine, by shuffling symbols as simple as "0" and "1", could simulate any conceivable form of mathematical reasoning.^{[338][339]} This, along with concurrent discoveries in [cybernetics](#), [information theory](#) and [neurobiology](#), led researchers to consider the possibility of building an "electronic brain".^[r] They developed several areas of research that would become part of AI,^[341] such as [McCulloch and Pitts](#) design for "artificial neurons" in 1943,^[115] and Turing's influential 1950 paper '[Computing Machinery and Intelligence](#)', which introduced the [Turing test](#) and showed that "machine intelligence" was plausible.^{[342][339]}

The field of AI research was founded at [a workshop](#) at [Dartmouth College](#) in 1956.^{[s][6]} The attendees became the leaders of AI research in the 1960s.^[t] They and their students produced programs that the press described as "astonishing":^[u] computers were learning [checkers](#) strategies, solving word problems in algebra, proving [logical theorems](#) and speaking English.^{[v][7]} Artificial intelligence laboratories were set up at a number of British and U.S. universities in the latter 1950s and early 1960s.^[339]

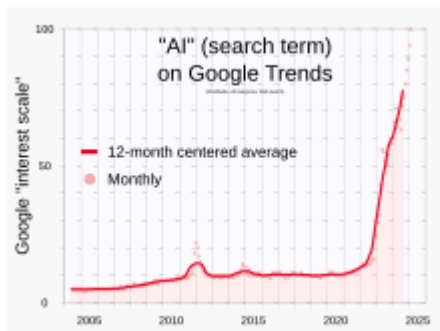
Researchers in the 1960s and the 1970s were convinced that their methods would eventually succeed in creating a machine with [general intelligence](#) and considered this the goal of their field.^[346] In 1965 [Herbert Simon](#) predicted, "machines will be capable, within twenty years, of doing any work a man can do".^[347] In 1967 [Marvin Minsky](#) agreed, writing that "within a generation ... the problem of creating 'artificial intelligence' will substantially be solved".^[348] They had, however, underestimated the difficulty of the problem.^[w] In 1974, both the U.S. and British governments cut off exploratory research in response to the [criticism](#) of [Sir James Lighthill](#)^[350] and ongoing pressure from the U.S. Congress to [fund more productive projects](#).^[351] [Minsky's](#) and [Papert's](#) book [Perceptrons](#) was understood as proving that [artificial neural networks](#) would never be useful for solving real-world tasks, thus discrediting the approach altogether.^[352] The "[AI winter](#)", a period when obtaining funding for AI projects was difficult, followed.^[9]

In the early 1980s, AI research was revived by the commercial success of [expert systems](#),^[353] a form of AI program that simulated the knowledge and analytical skills of human experts. By 1985, the market for AI had reached over a billion dollars. At the same time, Japan's [fifth generation computer](#) project inspired the U.S. and British governments to restore funding for [academic research](#).^[8] However, beginning with the collapse of the [Lisp Machine](#) market in 1987, AI once again fell into disrepute, and a second, longer-lasting winter began.^[10]

Up to this point, most of AI's funding had gone to projects that used high-level [symbols](#) to represent [mental objects](#) like plans, goals, beliefs, and known facts. In the 1980s, some researchers began to doubt that this approach would be able to imitate all the processes of human cognition, especially [perception](#), [robotics](#), [learning](#) and [pattern recognition](#),^[354] and began to look into "sub-symbolic" approaches.^[355] [Rodney Brooks](#) rejected "representation" in general and focussed directly on engineering machines that move and survive.^[x] [Judea Pearl](#), [Lofti Zadeh](#), and others developed methods that handled incomplete and uncertain information by making reasonable guesses rather than precise logic.^{[86][360]} But the most important development was the revival of "[connectionism](#)", including neural network research, by [Geoffrey Hinton](#) and others.^[361] In 1990, [Yann LeCun](#) successfully showed that [convolutional neural networks](#) can recognize handwritten digits, the first of many successful applications of neural networks.^[362]

AI gradually restored its reputation in the late 1990s and early 21st century by exploiting formal mathematical methods and by finding specific solutions to specific problems. This "[narrow](#)" and "formal" focus allowed researchers to produce verifiable results and collaborate with other fields (such as [statistics](#), [economics](#) and [mathematics](#)).^[363] By 2000, solutions developed by AI researchers were being widely used, although in the 1990s they were rarely described as "artificial intelligence" (a tendency known as the [AI effect](#)).^[364] However, several academic researchers became concerned that AI was no longer pursuing its original goal of creating versatile, fully intelligent machines. Beginning around 2002, they founded the subfield of [artificial general intelligence](#) (or "AGI"), which had several well-funded institutions by the 2010s.^[4]

[Deep learning](#) began to dominate industry benchmarks in 2012 and was adopted throughout the field.^[11] For many specific tasks, other methods were abandoned.^[v] Deep learning's success was based on both hardware improvements ([faster computers](#),^[366] [graphics processing units](#), [cloud computing](#)^[367]) and access to [large amounts of data](#)^[368] (including curated datasets,^[367] such as [ImageNet](#)). Deep learning's success led to an enormous increase in interest and funding in AI.^[2] The amount of machine learning research (measured by total publications) increased by 50% in the years 2015–2019.^[325]



The number of Google searches for the term "AI" began to increase in about 2022.

In 2016, issues of [fairness](#) and the misuse of technology were catapulted into center stage at machine learning conferences, publications vastly increased, funding became available, and many researchers re-focused their careers on these issues. The [alignment problem](#) became a serious field of academic study.^[302]

In the late 2010s and early 2020s, [AGI](#) companies began to deliver programs that created enormous interest. In 2015, [AlphaGo](#), developed by [DeepMind](#), beat the world champion [Go player](#). The program taught only the game's rules and developed a strategy by itself. [GPT-3](#) is a [large language model](#) that was released in 2020 by [OpenAI](#) and is capable of generating high-quality human-like text.^[369] [ChatGPT](#), launched on November 30, 2022, became the fastest-growing consumer software application in history, gaining over 100 million users in two months.^[370] It marked what is widely regarded as AI's breakout year, bringing it into the public consciousness.^[371] These programs, and others, inspired an aggressive [AI boom](#), where large companies began investing billions of dollars in AI research. According to AI Impacts, about \$50 billion annually was invested in "AI" around 2022 in the U.S. alone and about 20% of the new U.S. Computer Science PhD graduates have specialized in "AI".^[372] About 800,000 "AI"-related U.S. job openings existed in 2022.^[373] According to PitchBook research, 22% of newly funded [startups](#) in 2024 claimed to be AI companies.^[374]

Philosophy

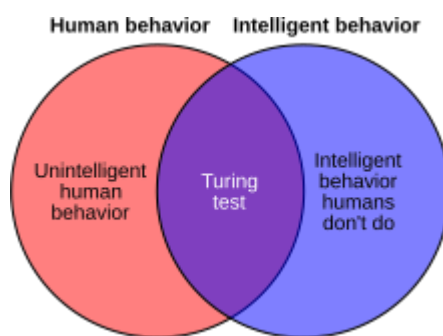
Main article: [Philosophy of artificial intelligence](#)

Philosophical debates have historically sought to determine the nature of intelligence and how to make intelligent machines.^[375] Another major focus has been whether machines can be conscious, and the associated ethical implications.^[376] Many other topics in philosophy are relevant to AI, such as [epistemology](#) and [free will](#).^[377] Rapid advancements have intensified public discussions on the philosophy and [ethics of AI](#).^[376]

Defining artificial intelligence

See also: [Turing test](#), [Intelligent agent](#), [Dartmouth workshop](#), and [Synthetic intelligence](#)

[Alan Turing](#) wrote in 1950 "I propose to consider the question 'can machines think?'"^[378] He advised changing the question from whether a machine "thinks", to "whether or not it is possible for machinery to show intelligent behaviour".^[378] He devised the Turing test, which measures the ability of a machine to simulate human conversation.^[342] Since we can only observe the behavior of the machine, it does not matter if it is "actually" thinking or literally has a "mind". Turing notes that [we can not determine these things about other people](#) but "it is usual to have a polite convention that everyone thinks."^[379]



The Turing test can provide some evidence of intelligence, but it penalizes non-human intelligent behavior.^[380]

[Russell](#) and [Norvig](#) agree with Turing that intelligence must be defined in terms of external behavior, not internal structure.^[1] However, they are critical that the test requires the machine to imitate humans. "[Aeronautical engineering](#) texts", they wrote, "do not define the goal of their field as making 'machines that fly so exactly like [pigeons](#) that they can fool other pigeons.'"^[381] AI founder [John McCarthy](#) agreed, writing that "Artificial intelligence is not, by definition, simulation of human intelligence".^[382]

McCarthy defines intelligence as "the computational part of the ability to achieve goals in the world".^[383] Another AI founder, [Marvin Minsky](#), similarly describes it as "the ability to solve hard problems".^[384] The leading AI textbook defines it as the study of agents that perceive their environment and take actions that maximize their chances of achieving defined goals.^[1] These definitions view intelligence in terms of well-defined problems with well-defined solutions, where both the difficulty of the problem and the performance of the program are direct measures of the "intelligence" of the machine—and no other philosophical discussion is required, or may not even be possible.

Another definition has been adopted by Google,^[385] a major practitioner in the field of AI. This definition stipulates the ability of systems to synthesize information as the manifestation of intelligence, similar to the way it is defined in biological intelligence.

Some authors have suggested in practice, that the definition of AI is vague and difficult to define, with contention as to whether classical algorithms should be categorised as AI,^[386] with many

companies during the early 2020s AI boom using the term as a marketing [buzzword](#), often even if they did "not actually use AI in a material way".^[387]

Evaluating approaches to AI

No established unifying theory or [paradigm](#) has guided AI research for most of its history.^[aa] The unprecedented success of statistical machine learning in the 2010s eclipsed all other approaches (so much so that some sources, especially in the business world, use the term "artificial intelligence" to mean "machine learning with neural networks"). This approach is mostly [sub-symbolic](#), [soft](#) and [narrow](#). Critics argue that these questions may have to be revisited by future generations of AI researchers.

Symbolic AI and its limits

[Symbolic AI](#) (or "[GOFAI](#)")^[389] simulated the high-level conscious reasoning that people use when they solve puzzles, express legal reasoning and do mathematics. They were highly successful at "intelligent" tasks such as algebra or IQ tests. In the 1960s, Newell and Simon proposed the [physical symbol systems hypothesis](#): "A physical symbol system has the necessary and sufficient means of general intelligent action."^[390]

However, the symbolic approach failed on many tasks that humans solve easily, such as learning, recognizing an object or commonsense reasoning. [Moravec's paradox](#) is the discovery that high-level "intelligent" tasks were easy for AI, but low level "instinctive" tasks were extremely difficult.^[391] Philosopher [Hubert Dreyfus](#) had [argued](#) since the 1960s that human expertise depends on unconscious instinct rather than conscious symbol manipulation, and on having a "feel" for the situation, rather than explicit symbolic knowledge.^[392] Although his arguments had been ridiculed and ignored when they were first presented, eventually, AI research came to agree with him.^{[ab][16]}

The issue is not resolved: [sub-symbolic](#) reasoning can make many of the same inscrutable mistakes that human intuition does, such as [algorithmic bias](#). Critics such as [Noam Chomsky](#) argue continuing research into symbolic AI will still be necessary to attain general intelligence,^{[394][395]} in part because sub-symbolic AI is a move away from [explainable AI](#): it can be difficult or impossible to understand why a modern statistical AI program made a particular decision. The emerging field of [neuro-symbolic artificial intelligence](#) attempts to bridge the two approaches.

Neat vs. scruffy

Main article: [Neats and scruffies](#)

"Neats" hope that intelligent behavior is described using simple, elegant principles (such as [logic](#), [optimization](#), or [neural networks](#)). "Scruffies" expect that it necessarily requires solving a large number of unrelated problems. Neats defend their programs with theoretical rigor, scruffies rely mainly on incremental testing to see if they work. This issue was actively discussed in the 1970s and 1980s,^[396] but eventually was seen as irrelevant. Modern AI has elements of both.

Soft vs. hard computing

Main article: [Soft computing](#)

Finding a provably correct or optimal solution is [intractable](#) for many important problems.^[15] Soft computing is a set of techniques, including [genetic algorithms](#), [fuzzy logic](#) and neural networks, that are tolerant of imprecision, uncertainty, partial truth and approximation. Soft computing was

introduced in the late 1980s and most successful AI programs in the 21st century are examples of soft computing with neural networks.

Narrow vs. general AI

Main articles: [Weak artificial intelligence](#) and [Artificial general intelligence](#)

AI researchers are divided as to whether to pursue the goals of artificial general intelligence and [superintelligence](#) directly or to solve as many specific problems as possible (narrow AI) in hopes these solutions will lead indirectly to the field's long-term goals.^{[397][398]} General intelligence is difficult to define and difficult to measure, and modern AI has had more verifiable successes by focusing on specific problems with specific solutions. The sub-field of artificial general intelligence studies this area exclusively.

Machine consciousness, sentience, and mind

Main articles: [Philosophy of artificial intelligence](#) and [Artificial consciousness](#)

The [philosophy of mind](#) does not know whether a machine can have a [mind](#), [consciousness](#) and [mental states](#), in the same sense that human beings do. This issue considers the internal experiences of the machine, rather than its external behavior. Mainstream AI research considers this issue irrelevant because it does not affect the goals of the field: to build machines that can solve problems using intelligence. [Russell](#) and [Norvig](#) add that "[t]he additional project of making a machine conscious in exactly the way humans are is not one that we are equipped to take on."^[399] However, the question has become central to the philosophy of mind. It is also typically the central question at issue in [artificial intelligence in fiction](#).

Consciousness

Main articles: [Hard problem of consciousness](#) and [Theory of mind](#)

[David Chalmers](#) identified two problems in understanding the mind, which he named the "hard" and "easy" problems of consciousness.^[400] The easy problem is understanding how the brain processes signals, makes plans and controls behavior. The hard problem is explaining how this *feels* or why it should feel like anything at all, assuming we are right in thinking that it truly does feel like something (Dennett's consciousness illusionism says this is an illusion). While human [information processing](#) is easy to explain, human [subjective experience](#) is difficult to explain. For example, it is easy to imagine a color-blind person who has learned to identify which objects in their field of view are red, but it is not clear what would be required for the person to *know what red looks like*.^[401]

Computationalism and functionalism

Main articles: [Computational theory of mind](#) and [Functionalism \(philosophy of mind\)](#)

Computationalism is the position in the [philosophy of mind](#) that the human mind is an information processing system and that thinking is a form of computing. Computationalism argues that the relationship between mind and body is similar or identical to the relationship between software and hardware and thus may be a solution to the [mind–body problem](#). This philosophical position was inspired by the work of AI researchers and cognitive scientists in the 1960s and was originally proposed by philosophers [Jerry Fodor](#) and [Hilary Putnam](#).^[402]

Philosopher [John Searle](#) characterized this position as "[strong AI](#)": "The appropriately programmed computer with the right inputs and outputs would thereby have a mind in exactly the same sense human beings have minds."^[403] Searle challenges this claim with his [Chinese room](#) argument, which

attempts to show that even a computer capable of perfectly simulating human behavior would not have a mind.^[406]

AI welfare and rights

It is difficult or impossible to reliably evaluate whether an advanced [AI is sentient](#) (has the ability to feel), and if so, to what degree.^[407] But if there is a significant chance that a given machine can feel and suffer, then it may be entitled to certain rights or welfare protection measures, similarly to animals.^{[408][409]} [Sapience](#) (a set of capacities related to high intelligence, such as discernment or [self-awareness](#)) may provide another moral basis for AI rights.^[408] [Robot rights](#) are also sometimes proposed as a practical way to integrate autonomous agents into society.^[410]

In 2017, the European Union considered granting "electronic personhood" to some of the most capable AI systems. Similarly to the legal status of companies, it would have conferred rights but also responsibilities.^[411] Critics argued in 2018 that granting rights to AI systems would downplay the importance of [human rights](#), and that legislation should focus on user needs rather than speculative futuristic scenarios. They also noted that robots lacked the autonomy to take part to society on their own.^{[412][413]}

Progress in AI increased interest in the topic. Proponents of AI welfare and rights often argue that AI sentience, if it emerges, would be particularly easy to deny. They warn that this may be a [moral blind spot](#) analogous to [slavery](#) or [factory farming](#), which could lead to [large-scale suffering](#) if sentient AI is created and carelessly exploited.^{[409][408]}

Future

Superintelligence and the singularity

A [superintelligence](#) is a hypothetical agent that would possess intelligence far surpassing that of the brightest and most gifted human mind.^[398] If research into [artificial general intelligence](#) produced sufficiently intelligent software, it might be able to [reprogram and improve itself](#). The improved software would be even better at improving itself, leading to what [I. J. Good](#) called an "[intelligence explosion](#)" and [Vernor Vinge](#) called a "[singularity](#)".^[414]

However, technologies cannot improve exponentially indefinitely, and typically follow an [S-shaped curve](#), slowing when they reach the physical limits of what the technology can do.^[415]

Transhumanism

Main article: [Transhumanism](#)

Robot designer [Hans Moravec](#), cyberneticist [Kevin Warwick](#) and inventor [Ray Kurzweil](#) have predicted that humans and machines may merge in the future into [cyborgs](#) that are more capable and powerful than either. This idea, called transhumanism, has roots in the writings of [Aldous Huxley](#) and [Robert Ettinger](#).^[416]

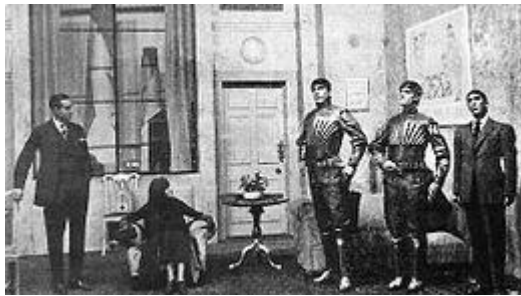
[Edward Fredkin](#) argues that "artificial intelligence is the next step in evolution", an idea first proposed by [Samuel Butler](#)'s "[Darwin among the Machines](#)" as far back as 1863, and expanded upon by [George Dyson](#) in his 1998 book [Darwin Among the Machines: The Evolution of Global Intelligence](#).^[417]

Decomputing

Arguments for *decomputing* have been raised by [Dan McQuillan](#) (*Resisting AI: An Anti-fascist Approach to Artificial Intelligence*, 2022), meaning an opposition to the sweeping application and expansion of artificial intelligence. Similar to [degrowth](#) the approach criticizes AI as an outgrowth of the systemic issues and capitalist world we live in. Arguing that a different future is possible, in which distance between people is reduced and not increased to AI intermediaries.^[418]

In fiction

Main article: [Artificial intelligence in fiction](#)



The word "robot" itself was coined by [Karel Čapek](#) in his 1921 play [R.U.R.](#), the title standing for "Rossum's Universal Robots".

Thought-capable artificial beings have appeared as storytelling devices since antiquity,^[419] and have been a persistent theme in [science fiction](#).^[420]

A common [trope](#) in these works began with [Mary Shelley's](#) *Frankenstein*, where a human creation becomes a threat to its masters. This includes such works as [Arthur C. Clarke's](#) and [Stanley Kubrick's](#) *2001: A Space Odyssey* (both 1968), with [HAL 9000](#), the murderous computer in charge of the *Discovery One* spaceship, as well as *The Terminator* (1984) and *The Matrix* (1999). In contrast, the rare loyal robots such as Gort from *The Day the Earth Stood Still* (1951) and Bishop from *Aliens* (1986) are less prominent in popular culture.^[421]

[Isaac Asimov](#) introduced the [Three Laws of Robotics](#) in many stories, most notably with the "Multivac" super-intelligent computer. Asimov's laws are often brought up during lay discussions of machine ethics;^[422] while almost all artificial intelligence researchers are familiar with Asimov's laws through popular culture, they generally consider the laws useless for many reasons, one of which is their ambiguity.^[423]

Several works use AI to force us to confront the fundamental question of what makes us human, showing us artificial beings that have [the ability to feel](#), and thus to suffer. This appears in [Karel Čapek's](#) *R.U.R.*, the films *A.I. Artificial Intelligence* and *Ex Machina*, as well as the novel *Do Androids Dream of Electric Sheep?*, by [Philip K. Dick](#). Dick considers the idea that our understanding of human subjectivity is altered by technology created with artificial intelligence.^[424]

See also

- [Artificial intelligence and elections](#) – Use and impact of AI on political elections
- [Artificial intelligence content detection](#) – Software to detect AI-generated content
- [Behavior selection algorithm](#) – Algorithm that selects actions for intelligent agents
- [Business process automation](#) – Automation of business processes
- [Case-based reasoning](#) – Process of solving new problems based on the solutions of similar past problems

- [Computational intelligence](#) – Ability of a computer to learn a specific task from data or experimental observation
- [Digital immortality](#) – Hypothetical concept of storing a personality in digital form
- [Emergent algorithm](#) – Algorithm exhibiting emergent behavior
- [Female gendering of AI technologies](#) – Gender biases in digital technology
- [Glossary of artificial intelligence](#) – List of definitions of terms and concepts commonly used in the study of artificial intelligence
- [Intelligence amplification](#) – Use of information technology to augment human intelligence
- [Intelligent agent](#) – Software agent which acts autonomously
- [Mind uploading](#) – Hypothetical process of digitally emulating a brain
- [Organoid intelligence](#) – Use of brain cells and brain organoids for intelligent computing
- [Robotic process automation](#) – Form of business process automation technology
- [The Last Day \(novel\)](#) – 1967 Welsh science fiction novel - Welsh science novel by [Owain Owain](#)
- [Wetware computer](#) – Computer composed of organic material

Explanatory notes

1. ^ [Jump up to:](#)^{a b} This list of intelligent traits is based on the topics covered by the major AI textbooks, including: [Russell & Norvig \(2021\)](#), [Luger & Stubblefield \(2004\)](#), [Poole, Mackworth & Goebel \(1998\)](#) and [Nilsson \(1998\)](#)
2. ^ [Jump up to:](#)^{a b} This list of tools is based on the topics covered by the major AI textbooks, including: [Russell & Norvig \(2021\)](#), [Luger & Stubblefield \(2004\)](#), [Poole, Mackworth & Goebel \(1998\)](#) and [Nilsson \(1998\)](#)
3. ^ It is among the reasons that [expert systems](#) proved to be inefficient for capturing knowledge.^{[30][31]}
4. ^ "Rational agent" is general term used in [economics](#), [philosophy](#) and theoretical artificial intelligence. It can refer to anything that directs its behavior to accomplish goals, such as a person, an animal, a corporation, a nation, or in the case of AI, a computer program.
5. ^ [Alan Turing](#) discussed the centrality of learning as early as 1950, in his classic paper "[Computing Machinery and Intelligence](#)".^[42] In 1956, at the original Dartmouth AI summer conference, [Ray Solomonoff](#) wrote a report on unsupervised probabilistic machine learning: "An Inductive Inference Machine".^[43]
6. ^ See [AI winter § Machine translation and the ALPAC report of 1966](#)
7. ^ Compared with symbolic logic, formal Bayesian inference is computationally expensive. For inference to be tractable, most observations must be [conditionally independent](#) of one another. [AdSense](#) uses a Bayesian network with over 300 million edges to learn which ads to serve.^[93]

8. [^] Expectation–maximization, one of the most popular algorithms in machine learning, allows clustering in the presence of unknown [latent variables](#).^[95]
9. [^] Some form of deep neural networks (without a specific learning algorithm) were described by: [Warren S. McCulloch](#) and [Walter Pitts](#) (1943)^[115] [Alan Turing](#) (1948);^[116] [Karl Steinbuch](#) and [Roger David Joseph](#) (1961).^[117] Deep or recurrent networks that learned (or used gradient descent) were developed by: [Frank Rosenblatt](#) (1957);^[116] [Oliver Selfridge](#) (1959);^[117] [Alexey Ivakhnenko](#) and [Valentin Lapa](#) (1965);^[118] [Kaoru Nakano](#) (1971);^[119] [Shun-Ichi Amari](#) (1972);^[119] [John Joseph Hopfield](#) (1982).^[119] Precursors to backpropagation were developed by: [Henry J. Kelley](#) (1960);^[116] [Arthur E. Bryson](#) (1962);^[116] [Stuart Dreyfus](#) (1962);^[116] [Arthur E. Bryson](#) and [Yu-Chi Ho](#) (1969);^[116] Backpropagation was independently developed by: [Seppo Linnainmaa](#) (1970);^[120] [Paul Werbos](#) (1974).^[116]
10. [^] [Geoffrey Hinton](#) said, of his work on neural networks in the 1990s, "our labeled datasets were thousands of times too small. [And] our computers were millions of times too slow."^[121]
11. [^] In statistics, a [bias](#) is a systematic error or deviation from the correct value. But in the context of [fairness](#), it refers to a tendency in favor or against a certain group or individual characteristic, usually in a way that is considered unfair or harmful. A statistically unbiased AI system that produces disparate outcomes for different demographic groups may thus be viewed as biased in the ethical sense.^[238]
12. [^] Including [Jon Kleinberg](#) ([Cornell University](#)), Sendhil Mullainathan ([University of Chicago](#)), Cynthia Chouldechova ([Carnegie Mellon](#)) and Sam Corbett-Davis ([Stanford](#))^[247]
13. [^] Moritz Hardt (a director at the [Max Planck Institute for Intelligent Systems](#)) argues that machine learning "is fundamentally the wrong tool for a lot of domains, where you're trying to design interventions and mechanisms that change the world."^[252]
14. [^] When the law was passed in 2018, it still contained a form of this provision.
15. [^] This is the [United Nations'](#) definition, and includes things like [land mines](#) as well.^[266]
16. [^] See table 4; 9% is both the OECD average and the U.S. average.^[277]
17. [^] Sometimes called a "[robopocalypse](#)"^[285]
18. [^] "Electronic brain" was the term used by the press around this time.^{[338][340]}
19. [^] Daniel Crevier wrote, "the conference is generally recognized as the official birthdate of the new science."^[343] [Russell](#) and [Norvig](#) called the conference "the inception of artificial intelligence."^[115]
20. [^] [Russell](#) and [Norvig](#) wrote "for the next 20 years the field would be dominated by these people and their students."^[344]
21. [^] [Russell](#) and [Norvig](#) wrote, "it was astonishing whenever a computer did anything kind of smartish".^[345]

22. [^] The programs described are [Arthur Samuel](#)'s checkers program for the [IBM 701](#), [Daniel Bobrow](#)'s [STUDENT](#), [Newell](#) and [Simon](#)'s [Logic Theorist](#) and [Terry Winograd](#)'s [SHRDLU](#).
23. [^] [Russell](#) and [Norvig](#) write: "in almost all cases, these early systems failed on more difficult problems"^[349]
24. [^] [Embodied](#) approaches to AI^[356] were championed by [Hans Moravec](#)^[357] and [Rodney Brooks](#)^[358] and went by many names: [Nouvelle AI](#).^[358] [Developmental robotics](#).^[359]
25. [^] Matteo Wong wrote in [The Atlantic](#): "Whereas for decades, computer-science fields such as natural-language processing, computer vision, and robotics used extremely different methods, now they all use a programming method called "deep learning". As a result, their code and approaches have become more similar, and their models are easier to integrate into one another."^[365]
26. [^] Jack Clark wrote in [Bloomberg](#): "After a half-decade of quiet breakthroughs in artificial intelligence, 2015 has been a landmark year. Computers are smarter and learning faster than ever", and noted that the number of software projects that use machine learning at [Google](#) increased from a "sporadic usage" in 2012 to more than 2,700 projects in 2015.^[367]
27. [^] [Nils Nilsson](#) wrote in 1983: "Simply put, there is wide disagreement in the field about what AI is all about."^[388]
28. [^] Daniel Crevier wrote that "time has proven the accuracy and perceptiveness of some of Dreyfus's comments. Had he formulated them less aggressively, constructive actions they suggested might have been taken much earlier."^[393]
29. [^] Searle presented this definition of "Strong AI" in 1999.^[403] Searle's original formulation was "The appropriately programmed computer really is a mind, in the sense that computers given the right programs can be literally said to understand and have other cognitive states."^[404] Strong AI is defined similarly by [Russell](#) and [Norvig](#): "Strong AI – the assertion that machines that do so are *actually* thinking (as opposed to *simulating* thinking)."^[405]

References

1. [^] [Jump up to:](#)^{a b c} [Russell & Norvig \(2021\)](#), pp. 1–4.
2. [^] [AI set to exceed human brain power Archived](#) 2008-02-19 at the [Wayback Machine](#) CNN.com (July 26, 2006)
3. [^] [Kaplan, Andreas; Haenlein, Michael \(2019\). "Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence". *Business Horizons*. **62**: 15–25. doi:10.1016/j.bushor.2018.08.004. ISSN 0007-6813. S2CID 158433736.](#)
4. [^] [Jump up to:](#)^{a b c} [Artificial general intelligence: Russell & Norvig \(2021\)](#), pp. 32–33, 1020–1021)
Proposal for the modern version: [Pennachin & Goertzel \(2007\)](#)
Warnings of overspecialization in AI from leading researchers: [Nilsson \(1995\)](#), [McCarthy \(2007\)](#), [Beal & Winston \(2009\)](#)

5. [^ Russell & Norvig \(2021, §1.2\).](#)
6. [^ Jump up to:^{a b} Dartmouth workshop: Russell & Norvig \(2021, p. 18\), McCorduck \(2004, pp. 111–136\), NRC \(1999, pp. 200–201\)](#)
The proposal: [McCarthy et al. \(1955\)](#)
7. [^ Jump up to:^{a b} Successful programs of the 1960s: McCorduck \(2004, pp. 243–252\), Crevier \(1993, pp. 52–107\), Moravec \(1988, p. 9\), Russell & Norvig \(2021, pp. 19–21\)](#)
8. [^ Jump up to:^{a b} Funding initiatives in the early 1980s: Fifth Generation Project \(Japan\), Alvey \(UK\), Microelectronics and Computer Technology Corporation \(US\), Strategic Computing Initiative \(US\): McCorduck \(2004, pp. 426–441\), Crevier \(1993, pp. 161–162, 197–203, 211, 240\), Russell & Norvig \(2021, p. 23\), NRC \(1999, pp. 210–211\), Newquist \(1994, pp. 235–248\)](#)
9. [^ Jump up to:^{a b} First AI Winter, Lighthill report, Mansfield Amendment: Crevier \(1993, pp. 115–117\), Russell & Norvig \(2021, pp. 21–22\), NRC \(1999, pp. 212–213\), Howe \(1994\), Newquist \(1994, pp. 189–201\)](#)
10. [^ Jump up to:^{a b} Second AI Winter: Russell & Norvig \(2021, p. 24\), McCorduck \(2004, pp. 430–435\), Crevier \(1993, pp. 209–210\), NRC \(1999, pp. 214–216\), Newquist \(1994, pp. 301–318\)](#)
11. [^ Jump up to:^{a b} Deep learning revolution, AlexNet: Goldman \(2022\), Russell & Norvig \(2021, p. 26\), McKinsey \(2018\)](#)
12. [^ Toews \(2023\).](#)
13. [^ Problem-solving, puzzle solving, game playing, and deduction: Russell & Norvig \(2021, chpt. 3–5\), Russell & Norvig \(2021, chpt. 6\) \(constraint satisfaction\), Poole, Mackworth & Goebel \(1998, chpt. 2, 3, 7, 9\), Luger & Stubblefield \(2004, chpt. 3, 4, 6, 8\), Nilsson \(1998, chpt. 7–12\)](#)
14. [^ Uncertain reasoning: Russell & Norvig \(2021, chpt. 12–18\), Poole, Mackworth & Goebel \(1998, pp. 345–395\), Luger & Stubblefield \(2004, pp. 333–381\), Nilsson \(1998, chpt. 7–12\)](#)
15. [^ Jump up to:^{a b c} Intractability and efficiency and the combinatorial explosion: Russell & Norvig \(2021, p. 21\)](#)
16. [^ Jump up to:^{a b c} Psychological evidence of the prevalence of sub-symbolic reasoning and knowledge: Kahneman \(2011\), Dreyfus & Dreyfus \(1986\), Wason & Shapiro \(1966\), Kahneman, Slovic & Tversky \(1982\)](#)
17. [^ Knowledge representation and knowledge engineering: Russell & Norvig \(2021, chpt. 10\), Poole, Mackworth & Goebel \(1998, pp. 23–46, 69–81, 169–233, 235–277, 281–298, 319–345\), Luger & Stubblefield \(2004, pp. 227–243\), Nilsson \(1998, chpt. 17.1–17.4, 18\)](#)
18. [^ Smoliar & Zhang \(1994\).](#)
19. [^ Neumann & Möller \(2008\).](#)

20. [^ Kuperman, Reichley & Bailey \(2006\).](#)
21. [^ McGarry \(2005\).](#)
22. [^ Bertini, Del Bimbo & Torniai \(2006\).](#)
23. [^ Russell & Norvig \(2021\),](#) pp. 272.
24. [^](#) Representing categories and relations: [Semantic networks](#), [description logics](#), [inheritance](#) (including [frames](#), and [scripts](#)): [Russell & Norvig \(2021, §10.2 & 10.5\)](#), [Poole, Mackworth & Goebel \(1998, pp. 174–177\)](#), [Luger & Stubblefield \(2004, pp. 248–258\)](#), [Nilsson \(1998, chpt. 18.3\)](#)
25. [^](#) Representing events and time: [Situation calculus](#), [event calculus](#), [fluent calculus](#) (including solving the [frame problem](#)): [Russell & Norvig \(2021, §10.3\)](#), [Poole, Mackworth & Goebel \(1998, pp. 281–298\)](#), [Nilsson \(1998, chpt. 18.2\)](#)
26. [^](#) [Causal calculus](#): [Poole, Mackworth & Goebel \(1998, pp. 335–337\)](#)
27. [^](#) Representing knowledge about knowledge: Belief calculus, [modal logics](#): [Russell & Norvig \(2021, §10.4\)](#), [Poole, Mackworth & Goebel \(1998, pp. 275–277\)](#)
28. [^](#) [Jump up to:](#)^a [^](#) [Default reasoning](#), [Frame problem](#), [default logic](#), [non-monotonic logics](#), [circumscription](#), [closed world assumption](#), [abduction](#): [Russell & Norvig \(2021, §10.6\)](#), [Poole, Mackworth & Goebel \(1998, pp. 248–256, 323–335\)](#), [Luger & Stubblefield \(2004, pp. 335–363\)](#), [Nilsson \(1998, ~18.3.3\)](#) (Poole *et al.* places abduction under "default reasoning". Luger *et al.* places this under "uncertain reasoning").
29. [^](#) [Jump up to:](#)^a [^](#) Breadth of commonsense knowledge: [Lenat & Guha \(1989, Introduction\)](#), [Crevier \(1993, pp. 113–114\)](#), [Moravec \(1988, p. 13\)](#), [Russell & Norvig \(2021, pp. 241, 385, 982\)](#) ([qualification problem](#))
30. [^](#) [Newquist \(1994\),](#) p. 296.
31. [^](#) [Crevier \(1993\),](#) pp. 204–208.
32. [^](#) [Russell & Norvig \(2021\),](#) p. 528.
33. [^](#) [Automated planning](#): [Russell & Norvig \(2021, chpt. 11\).](#)
34. [^](#) [Automated decision making](#), [Decision theory](#): [Russell & Norvig \(2021, chpt. 16–18\).](#)
35. [^](#) [Classical planning](#): [Russell & Norvig \(2021, Section 11.2\).](#)
36. [^](#) Sensorless or "conformant" planning, contingent planning, replanning (a.k.a online planning): [Russell & Norvig \(2021, Section 11.5\).](#)
37. [^](#) Uncertain preferences: [Russell & Norvig \(2021, Section 16.7\)](#) [Inverse reinforcement learning](#): [Russell & Norvig \(2021, Section 22.6\)](#)
38. [^](#) [Information value theory](#): [Russell & Norvig \(2021, Section 16.6\).](#)
39. [^](#) [Markov decision process](#): [Russell & Norvig \(2021, chpt. 17\).](#)
40. [^](#) [Game theory](#) and multi-agent decision theory: [Russell & Norvig \(2021, chpt. 18\).](#)

41. [^ Learning](#): [Russell & Norvig \(2021, chpt. 19–22\)](#), [Poole, Mackworth & Goebel \(1998, pp. 397–438\)](#), [Luger & Stubblefield \(2004, pp. 385–542\)](#), [Nilsson \(1998, chpt. 3.3, 10.3, 17.5, 20\)](#)
42. [^ Turing \(1950\)](#).
43. [^ Solomonoff \(1956\)](#).
44. [^ Unsupervised learning](#): [Russell & Norvig \(2021, pp. 653\)](#) (definition), [Russell & Norvig \(2021, pp. 738–740\)](#) (cluster analysis), [Russell & Norvig \(2021, pp. 846–860\)](#) (word embedding)
45. [^ Jump up to:](#)^a [^ Supervised learning](#): [Russell & Norvig \(2021, §19.2\)](#) (Definition), [Russell & Norvig \(2021, Chpt. 19–20\)](#) (Techniques)
46. [^ Reinforcement learning](#): [Russell & Norvig \(2021, chpt. 22\)](#), [Luger & Stubblefield \(2004, pp. 442–449\)](#)
47. [^ Transfer learning](#): [Russell & Norvig \(2021, pp. 281\)](#), [The Economist \(2016\)](#)
48. [^ "Artificial Intelligence \(AI\): What Is AI and How Does It Work? | Built In"](#). [builtin.com](#). Retrieved 30 October 2023.
49. [^ Computational learning theory](#): [Russell & Norvig \(2021, pp. 672–674\)](#), [Jordan & Mitchell \(2015\)](#)
50. [^ Natural language processing \(NLP\)](#): [Russell & Norvig \(2021, chpt. 23–24\)](#), [Poole, Mackworth & Goebel \(1998, pp. 91–104\)](#), [Luger & Stubblefield \(2004, pp. 591–632\)](#)
51. [^ Subproblems of NLP](#): [Russell & Norvig \(2021, pp. 849–850\)](#)
52. [^ Russell & Norvig \(2021\)](#), pp. 856–858.
53. [^ Dickson \(2022\)](#).
54. [^ Modern statistical and deep learning approaches to NLP](#): [Russell & Norvig \(2021, chpt. 24\)](#), [Cambria & White \(2014\)](#)
55. [^ Vincent \(2019\)](#).
56. [^ Russell & Norvig \(2021\)](#), pp. 875–878.
57. [^ Bushwick \(2023\)](#).
58. [^ Computer vision](#): [Russell & Norvig \(2021, chpt. 25\)](#), [Nilsson \(1998, chpt. 6\)](#)
59. [^ Russell & Norvig \(2021\)](#), pp. 849–850.
60. [^ Russell & Norvig \(2021\)](#), pp. 895–899.
61. [^ Russell & Norvig \(2021\)](#), pp. 899–901.
62. [^ Challa et al. \(2011\)](#).
63. [^ Russell & Norvig \(2021\)](#), pp. 931–938.
64. [^ MIT AIL \(2014\)](#).

65. [^ Affective computing](#): [Thro \(1993\)](#), [Edelson \(1991\)](#), [Tao & Tan \(2005\)](#), [Scassellati \(2002\)](#)
66. [^ Waddell \(2018\)](#).
67. [^ Poria et al. \(2017\)](#).
68. [^ Search algorithms](#): [Russell & Norvig \(2021\)](#), chpts. 3–5), [Poole, Mackworth & Goebel \(1998\)](#), pp. 113–163), [Luger & Stubblefield \(2004\)](#), pp. 79–164, 193–219), [Nilsson \(1998\)](#), chpts. 7–12)
69. [^ State space search](#): [Russell & Norvig \(2021\)](#), chpt. 3)
70. [^ Russell & Norvig \(2021\)](#), sect. 11.2.
71. [^ Uninformed searches](#) ([breadth first search](#), [depth-first search](#) and general [state space search](#)): [Russell & Norvig \(2021\)](#), sect. 3.4), [Poole, Mackworth & Goebel \(1998\)](#), pp. 113–132), [Luger & Stubblefield \(2004\)](#), pp. 79–121), [Nilsson \(1998\)](#), chpt. 8)
72. [^ Heuristic](#) or informed searches (e.g., greedy [best first](#) and [A*](#)): [Russell & Norvig \(2021\)](#), sect. 3.5), [Poole, Mackworth & Goebel \(1998\)](#), pp. 132–147), [Poole & Mackworth \(2017\)](#), sect. 3.6), [Luger & Stubblefield \(2004\)](#), pp. 133–150)
73. [^ Adversarial search](#): [Russell & Norvig \(2021\)](#), chpt. 5)
74. [^ Local](#) or "[optimization](#)" search: [Russell & Norvig \(2021\)](#), chpt. 4)
75. [^ Singh Chauhan, Nagesh \(18 December 2020\)](#). "[Optimization Algorithms in Neural Networks](#)". *KDnuggets*. Retrieved 13 January 2024.
76. [^ Evolutionary computation](#): [Russell & Norvig \(2021\)](#), sect. 4.1.2)
77. [^ Merkle & Middendorf \(2013\)](#).
78. [^ Logic](#): [Russell & Norvig \(2021\)](#), chpts. 6–9), [Luger & Stubblefield \(2004\)](#), pp. 35–77), [Nilsson \(1998\)](#), chpt. 13–16)
79. [^ Propositional logic](#): [Russell & Norvig \(2021\)](#), chpt. 6), [Luger & Stubblefield \(2004\)](#), pp. 45–50), [Nilsson \(1998\)](#), chpt. 13)
80. [^ First-order logic](#) and features such as [equality](#): [Russell & Norvig \(2021\)](#), chpt. 7), [Poole, Mackworth & Goebel \(1998\)](#), pp. 268–275), [Luger & Stubblefield \(2004\)](#), pp. 50–62), [Nilsson \(1998\)](#), chpt. 15)
81. [^ Logical inference](#): [Russell & Norvig \(2021\)](#), chpt. 10)
82. [^ logical deduction as search](#): [Russell & Norvig \(2021\)](#), sects. 9.3, 9.4), [Poole, Mackworth & Goebel \(1998\)](#), pp. ~46–52), [Luger & Stubblefield \(2004\)](#), pp. 62–73), [Nilsson \(1998\)](#), chpt. 4.2, 7.2)
83. [^ Resolution](#) and [unification](#): [Russell & Norvig \(2021\)](#), sections 7.5.2, 9.2, 9.5)
84. [^ Warren, D.H.; Pereira, L.M.; Pereira, F. \(1977\)](#). "Prolog-the language and its implementation compared with Lisp". *ACM SIGPLAN Notices*. **12** (8): 109–115. doi:10.1145/872734.806939.
85. [^ Fuzzy logic](#): [Russell & Norvig \(2021\)](#), pp. 214, 255, 459), [Scientific American \(1999\)](#)

86. ^ [Jump up to:](#)^{a b} Stochastic methods for uncertain reasoning: [Russell & Norvig \(2021, chpt. 12–18, 20\)](#), [Poole, Mackworth & Goebel \(1998, pp. 345–395\)](#), [Luger & Stubblefield \(2004, pp. 165–191, 333–381\)](#), [Nilsson \(1998, chpt. 19\)](#)
87. ^ [decision theory](#) and [decision analysis](#): [Russell & Norvig \(2021, chpt. 16–18\)](#), [Poole, Mackworth & Goebel \(1998, pp. 381–394\)](#)
88. ^ [Information value theory](#): [Russell & Norvig \(2021, sect. 16.6\)](#)
89. ^ [Markov decision processes](#) and dynamic [decision networks](#): [Russell & Norvig \(2021, chpt. 17\)](#)
90. ^ [Jump up to:](#)^{a b c} Stochastic temporal models: [Russell & Norvig \(2021, chpt. 14\)](#) [Hidden Markov model](#): [Russell & Norvig \(2021, sect. 14.3\)](#) [Kalman filters](#): [Russell & Norvig \(2021, sect. 14.4\)](#) [Dynamic Bayesian networks](#): [Russell & Norvig \(2021, sect. 14.5\)](#)
91. ^ [Game theory](#) and [mechanism design](#): [Russell & Norvig \(2021, chpt. 18\)](#)
92. ^ [Bayesian networks](#): [Russell & Norvig \(2021, sects. 12.5–12.6, 13.4–13.5, 14.3–14.5, 16.5, 20.2–20.3\)](#), [Poole, Mackworth & Goebel \(1998, pp. 361–381\)](#), [Luger & Stubblefield \(2004, pp. ~182–190, ~363–379\)](#), [Nilsson \(1998, chpt. 19.3–19.4\)](#)
93. ^ [Domingos \(2015\)](#), chpt. 6.
94. ^ [Bayesian inference](#) algorithm: [Russell & Norvig \(2021, sect. 13.3–13.5\)](#), [Poole, Mackworth & Goebel \(1998, pp. 361–381\)](#), [Luger & Stubblefield \(2004, pp. ~363–379\)](#), [Nilsson \(1998, chpt. 19.4 & 7\)](#)
95. ^ [Domingos \(2015\)](#), p. 210.
96. ^ [Bayesian learning](#) and the [expectation–maximization algorithm](#): [Russell & Norvig \(2021, chpt. 20\)](#), [Poole, Mackworth & Goebel \(1998, pp. 424–433\)](#), [Nilsson \(1998, chpt. 20\)](#), [Domingos \(2015, p. 210\)](#)
97. ^ [Bayesian decision theory](#) and Bayesian [decision networks](#): [Russell & Norvig \(2021, sect. 16.5\)](#)
98. ^ Statistical learning methods and [classifiers](#): [Russell & Norvig \(2021, chpt. 20\)](#),
99. ^ [Ciaramella, Alberto](#); Ciaramella, Marco (2024). *Introduction to Artificial Intelligence: from data analysis to generative AI*. Intellisemantic Editions. [ISBN 978-8-8947-8760-3](#).
100. ^ [Decision trees](#): [Russell & Norvig \(2021, sect. 19.3\)](#), [Domingos \(2015, p. 88\)](#)
101. ^ [Non-parameteric](#) learning models such as [K-nearest neighbor](#) and [support vector machines](#): [Russell & Norvig \(2021, sect. 19.7\)](#), [Domingos \(2015, p. 187\)](#) (k-nearest neighbor)
 - [Domingos \(2015, p. 88\)](#) (kernel methods)
102. ^ [Domingos \(2015\)](#), p. 152.
103. ^ [Naive Bayes classifier](#): [Russell & Norvig \(2021, sect. 12.6\)](#), [Domingos \(2015, p. 152\)](#)

104. [^] [Jump up to:^{a b} Neural networks: Russell & Norvig \(2021, chpt. 21\), Domingos \(2015, Chapter 4\)](#)
105. [^] [Gradient calculation in computational graphs, backpropagation, automatic differentiation: Russell & Norvig \(2021, sect. 21.2\), Luger & Stubblefield \(2004, pp. 467–474\), Nilsson \(1998, chpt. 3.3\)](#)
106. [^] [Universal approximation theorem: Russell & Norvig \(2021, p. 752\) The theorem: Cybenko \(1988\), Hornik, Stinchcombe & White \(1989\)](#)
107. [^] [Feedforward neural networks: Russell & Norvig \(2021, sect. 21.1\)](#)
108. [^] [Recurrent neural networks: Russell & Norvig \(2021, sect. 21.6\)](#)
109. [^] [Perceptrons: Russell & Norvig \(2021, pp. 21, 22, 683, 22\)](#)
110. [^] [Jump up to:^{a b} Deep learning: Russell & Norvig \(2021, chpt. 21\), Goodfellow, Bengio & Courville \(2016\), Hinton et al. \(2016\), Schmidhuber \(2015\)](#)
111. [^] [Convolutional neural networks: Russell & Norvig \(2021, sect. 21.3\)](#)
112. [^] [Deng & Yu \(2014\), pp. 199–200.](#)
113. [^] [Ciresan, Meier & Schmidhuber \(2012\).](#)
114. [^] [Russell & Norvig \(2021\), p. 750.](#)
115. [^] [Jump up to:^{a b c} Russell & Norvig \(2021\), p. 17.](#)
116. [^] [Jump up to:^{a b c d e f g} Russell & Norvig \(2021\), p. 785.](#)
117. [^] [Jump up to:^{a b} Schmidhuber \(2022\), sect. 5.](#)
118. [^] [Schmidhuber \(2022\), sect. 6.](#)
119. [^] [Jump up to:^{a b c} Schmidhuber \(2022\), sect. 7.](#)
120. [^] [Schmidhuber \(2022\), sect. 8.](#)
121. [^] [Quoted in Christian \(2020, p. 22\)](#)
122. [^] [Smith \(2023\).](#)
123. [^] ["Explained: Generative AI". 9 November 2023.](#)
124. [^] ["AI Writing and Content Creation Tools". MIT Sloan Teaching & Learning Technologies. Archived from the original on 25 December 2023. Retrieved 25 December 2023.](#)
125. [^] [Marmouyet \(2023\).](#)
126. [^] [Kobielus \(2019\).](#)
127. [^] [Thomason, James \(21 May 2024\). "Mojo Rising: The resurgence of AI-first programming languages". VentureBeat. Archived from the original on 27 June 2024. Retrieved 26 May 2024.](#)

128. [^] Wodecki, Ben (5 May 2023). ["7 AI Programming Languages You Need to Know"](#). AI Business. [Archived](#) from the original on 25 July 2024. Retrieved 5 October 2024.
129. [^] Plumb, Taryn (18 September 2024). ["Why Jensen Huang and Marc Benioff see 'gigantic' opportunity for agentic AI"](#). VentureBeat. [Archived](#) from the original on 5 October 2024. Retrieved 4 October 2024.
130. [^] Mims, Christopher (19 September 2020). ["Huang's Law Is the New Moore's Law, and Explains Why Nvidia Wants Arm"](#). Wall Street Journal. [ISSN 0099-9660](#). [Archived](#) from the original on 2 October 2023. Retrieved 19 January 2025.
131. [^] Davenport, T; Kalakota, R (June 2019). ["The potential for artificial intelligence in healthcare"](#). *Future Healthc J.* **6** (2): 94–98. [doi:10.7861/futurehosp.6-2-94](#). [PMC 6616181](#). [PMID 31363513](#).
132. [^] Lyakhova, U.A.; Lyakhov, P.A. (2024). ["Systematic review of approaches to detection and classification of skin cancer using artificial intelligence: Development and prospects"](#). *Computers in Biology and Medicine.* **178**: 108742. [doi:10.1016/j.combiomed.2024.108742](#). [PMID 38875908](#). [Archived](#) from the original on 3 December 2024. Retrieved 10 October 2024.
133. [^] Alqudaihi, Kawther S.; Aslam, Nida; Khan, Irfan Ullah; Almuhaideb, Abdullah M.; Alsunaidi, Shikah J.; Ibrahim, Nehad M. Abdel Rahman; Alhaidari, Fahd A.; Shaikh, Fatema S.; Alsenbel, Yasmine M.; Alalharith, Dima M.; Alharthi, Hajar M.; Alghamdi, Wejdan M.; Alshahrani, Mohammed S. (2021). ["Cough Sound Detection and Diagnosis Using Artificial Intelligence Techniques: Challenges and Opportunities"](#). *IEEE Access.* **9**: 102327–102344. [Bibcode:2021IEEEA...9j2327A](#). [doi:10.1109/ACCESS.2021.3097559](#). [ISSN 2169-3536](#). [PMC 8545201](#). [PMID 34786317](#).
134. [^] ^a ^b [Jump up to:](#) Bax, Monique; Thorpe, Jordan; Romanov, Valentin (December 2023). ["The future of personalized cardiovascular medicine demands 3D and 4D printing, stem cells, and artificial intelligence"](#). *Frontiers in Sensors.* **4**. [doi:10.3389/fsens.2023.1294721](#). [ISSN 2673-5067](#).
135. [^] Jumper, J; Evans, R; Pritzel, A (2021). ["Highly accurate protein structure prediction with AlphaFold"](#). *Nature.* **596** (7873): 583–589. [Bibcode:2021Natur.596..583J](#). [doi:10.1038/s41586-021-03819-2](#). [PMC 8371605](#). [PMID 34265844](#).
136. [^] ["AI discovers new class of antibiotics to kill drug-resistant bacteria"](#). 20 December 2023. [Archived](#) from the original on 16 September 2024. Retrieved 5 October 2024.
137. [^] ["AI speeds up drug design for Parkinson's ten-fold"](#). Cambridge University. 17 April 2024. [Archived](#) from the original on 5 October 2024. Retrieved 5 October 2024.
138. [^] Horne, Robert I.; Andrzejewska, Ewa A.; Alam, Parvez; Brotzakis, Z. Faidon; Srivastava, Ankit; Aubert, Alice; Nowinska, Magdalena; Gregory, Rebecca C.; Staats, Roxine; Possenti, Andrea; Chia, Sean; Sormanni, Pietro; Ghetti, Bernardino; Caughey,

- Byron; Knowles, Tuomas P. J.; Vendruscolo, Michele (17 April 2024). ["Discovery of potent inhibitors of \$\alpha\$ -synuclein aggregation using structure-based iterative learning"](#). *Nature Chemical Biology*. **20** (5). Nature: 634–645. doi:[10.1038/s41589-024-01580-x](#). PMC [11062903](#). PMID [38632492](#).
139. [^](#) Grant, Eugene F.; Lardner, Rex (25 July 1952). ["The Talk of the Town – It"](#). *The New Yorker*. ISSN [0028-792X](#). Archived from the original on 16 February 2020. Retrieved 28 January 2024.
 140. [^](#) Anderson, Mark Robert (11 May 2017). ["Twenty years on from Deep Blue vs Kasparov: how a chess match started the big data revolution"](#). *The Conversation*. Archived from the original on 17 September 2024. Retrieved 28 January 2024.
 141. [^](#) Markoff, John (16 February 2011). ["Computer Wins on 'Jeopardy!': Trivial, It's Not"](#). *The New York Times*. ISSN [0362-4331](#). Archived from the original on 22 October 2014. Retrieved 28 January 2024.
 142. [^](#) Byford, Sam (27 May 2017). ["AlphaGo retires from competitive Go after defeating world number one 3–0"](#). *The Verge*. Archived from the original on 7 June 2017. Retrieved 28 January 2024.
 143. [^](#) Brown, Noam; Sandholm, Tuomas (30 August 2019). ["Superhuman AI for multiplayer poker"](#). *Science*. **365** (6456): 885–890. Bibcode:[2019Sci...365..885B](#). doi:[10.1126/science.aay2400](#). ISSN [0036-8075](#). PMID [31296650](#).
 144. [^](#) ["MuZero: Mastering Go, chess, shogi and Atari without rules"](#). Google DeepMind. 23 December 2020. Retrieved 28 January 2024.
 145. [^](#) Sample, Ian (30 October 2019). ["AI becomes grandmaster in 'fiendishly complex' StarCraft II"](#). *The Guardian*. ISSN [0261-3077](#). Archived from the original on 29 December 2020. Retrieved 28 January 2024.
 146. [^](#) Wurman, P. R.; Barrett, S.; Kawamoto, K. (2022). ["Outracing champion Gran Turismo drivers with deep reinforcement learning"](#) (PDF). *Nature*. **602** (7896): 223–228. Bibcode:[2022Natur.602..223W](#). doi:[10.1038/s41586-021-04357-7](#). PMID [35140384](#).
 147. [^](#) Wilkins, Alex (13 March 2024). ["Google AI learns to play open-world video games by watching them"](#). *New Scientist*. Archived from the original on 26 July 2024. Retrieved 21 July 2024.
 148. [^](#) Wu, Zhengxuan; Arora, Aryaman; Wang, Zheng; Geiger, Atticus; Jurafsky, Dan; Manning, Christopher D.; Potts, Christopher (2024). ["ReFT: Representation Finetuning for Language Models"](#). *NeurIPS*. arXiv:[2404.03592](#).
 149. [^](#) ["Improving mathematical reasoning with process supervision"](#). OpenAI. 31 May 2023. Retrieved 26 January 2025.

150. [^](#) Srivastava, Saurabh (29 February 2024). "Functional Benchmarks for Robust Evaluation of Reasoning Performance, and the Reasoning Gap". [arXiv:2402.19450 \[cs.AI\]](#).
151. [^](#) Lightman, Hunter; Kosaraju, Vineet; Burda, Yura; Edwards, Harri; Baker, Bowen; Lee, Teddy; Leike, Jan; Schulman, John; Sutskever, Ilya; Cobbe, Karl (2023). "Let's Verify Step by Step". [arXiv:2305.20050v1 \[cs.LG\]](#).
152. [^](#) Franzen, Carl (8 August 2024). "[Alibaba claims no. 1 spot in AI math models with Qwen2-Math](#)". VentureBeat. Retrieved 16 February 2025.
153. [^](#) Franzen, Carl (9 January 2025). "[Microsoft's new rStar-Math technique upgrades small models to outperform OpenAI's o1-preview at math problems](#)". VentureBeat. Retrieved 26 January 2025.
154. [^](#) Roberts, Siobhan (25 July 2024). "[AI achieves silver-medal standard solving International Mathematical Olympiad problems](#)". *The New York Times*. [Archived](#) from the original on 26 September 2024. Retrieved 7 August 2024.
155. [^](#) Azerbayev, Zhangir; Schoelkopf, Hailey; Paster, Keiran; Santos, Marco Dos; McAleer, Stephen; Jiang, Albert Q.; Deng, Jia; Biderman, Stella; Welleck, Sean (16 October 2023). "[Llemma: An Open Language Model For Mathematics](#)". EleutherAI Blog. Retrieved 26 January 2025.
156. [^](#) "[Julius AI](#)". [julius.ai](#).
157. [^](#) McFarland, Alex (12 July 2024). "[8 Best AI for Math Tools \(January 2025\)](#)". Unite.AI. Retrieved 26 January 2025.
158. [^](#) Matthew Finio & Amanda Downie: IBM Think 2024 Primer, "What is Artificial Intelligence (AI) in Finance?" 8 Dec. 2023
159. [^](#) M. Nicolas, J. Firzli: Pensions Age/European Pensions magazine, "Artificial Intelligence: Ask the Industry" May June 2024 <https://videovoice.org/ai-in-finance-innovation-entrepreneurship-vs-over-regulation-with-the-eus-artificial-intelligence-act-wont-work-as-intended/> [Archived](#) 11 September 2024 at the [Wayback Machine](#).
160. [^](#) [Jump up to:](#) ^a ^b ^c Congressional Research Service (2019). [Artificial Intelligence and National Security](#) (PDF). Washington, DC: Congressional Research Service. [Archived](#) (PDF) from the original on 8 May 2020. Retrieved 25 February 2024.[PD-notice](#)
161. [^](#) [Jump up to:](#) ^a ^b Slyusar, Vadym (2019). Artificial intelligence as the basis of future control networks (Preprint). [doi:10.13140/RG.2.2.30247.50087](#).
162. [^](#) Iraqi, Amjad (3 April 2024). "['Lavender': The AI machine directing Israel's bombing spree in Gaza](#)". +972 Magazine. [Archived](#) from the original on 10 October 2024. Retrieved 6 April 2024.
163. [^](#) Davies, Harry; McKernan, Bethan; Sabbagh, Dan (1 December 2023). "['The Gospel': how Israel uses AI to select bombing targets in Gaza](#)". *The Guardian*. [Archived](#) from the original on 6 December 2023. Retrieved 4 December 2023.

164. [^] Marti, J Werner (10 August 2024). ["Drohnen haben den Krieg in der Ukraine revolutioniert, doch sie sind empfindlich auf Störsender – deshalb sollen sie jetzt autonom operieren"](#). Neue Zürcher Zeitung (in German). [Archived](#) from the original on 10 August 2024. Retrieved 10 August 2024.
165. [^] Newsom, Gavin; Weber, Shirley N. (5 September 2023). ["Executive Order N-12-23"](#) (PDF). Executive Department, State of California. [Archived](#) (PDF) from the original on 21 February 2024. Retrieved 7 September 2023.
166. [^] Pinaya, Walter H. L.; Graham, Mark S.; Kerfoot, Eric; Tudosiu, Petru-Daniel; Dafflon, Jessica; Fernandez, Virginia; Sanchez, Pedro; Wolleb, Julia; da Costa, Pedro F.; Patel, Ashay (2023). "Generative AI for Medical Imaging: extending the MONAI Framework". [arXiv:2307.15208 \[eess.IV\]](#).
167. [^] ["What is ChatGPT, DALL-E, and generative AI?"](#). McKinsey. Retrieved 14 December 2024.
168. [^] ["What is generative AI?"](#). IBM. 22 March 2024.
169. [^] Pasick, Adam (27 March 2023). ["Artificial Intelligence Glossary: Neural Networks and Other Terms Explained"](#). The New York Times. [ISSN 0362-4331](#). [Archived](#) from the original on 1 September 2023. Retrieved 22 April 2023.
170. [^] Karpathy, Andrej; Abbeel, Pieter; Brockman, Greg; Chen, Peter; Cheung, Vicki; Duan, Yan; Goodfellow, Ian; Kingma, Durk; Ho, Jonathan; Rein Houthooft; Tim Salimans; John Schulman; Ilya Sutskever; Wojciech Zaremba (16 June 2016). ["Generative models"](#). OpenAI. [Archived](#) from the original on 17 November 2023. Retrieved 15 March 2023.
171. [^] ^a ^b [Jump up to:](#) Griffith, Erin; Metz, Cade (27 January 2023). ["Anthropic Said to Be Closing In on \\$300 Million in New A.I. Funding"](#). [The New York Times](#). [Archived](#) from the original on 9 December 2023. Retrieved 14 March 2023.
172. [^] Lanxon, Nate; Bass, Dina; Davalos, Jackie (10 March 2023). ["A Cheat Sheet to AI Buzzwords and Their Meanings"](#). Bloomberg News. [Archived](#) from the original on 17 November 2023. Retrieved 14 March 2023.
173. [^] Metz, Cade (14 March 2023). ["OpenAI Plans to Up the Ante in Tech's A.I. Race"](#). The New York Times. [ISSN 0362-4331](#). [Archived](#) from the original on 31 March 2023. Retrieved 31 March 2023.
174. [^] Thoppilan, Romal; De Freitas, Daniel; Hall, Jamie; Shazeer, Noam; Kulshreshtha, Apoorv (20 January 2022). "LaMDA: Language Models for Dialog Applications". [arXiv:2201.08239 \[cs.CL\]](#).
175. [^] Roose, Kevin (21 October 2022). ["A Coming-Out Party for Generative A.I., Silicon Valley's New Craze"](#). [The New York Times](#). [Archived](#) from the original on 15 February 2023. Retrieved 14 March 2023.
176. [^] Metz, Cade (15 February 2024). ["OpenAI Unveils A.I. That Instantly Generates Eye-Popping Videos"](#). The New York Times. [ISSN 0362-4331](#). [Archived](#) from the original on 15 February 2024. Retrieved 16 February 2024.

177. [^ "The race of the AI labs heats up"](#). *The Economist*. 30 January 2023. [Archived](#) from the original on 17 November 2023. Retrieved 14 March 2023.
178. [^](#) Yang, June; Gokturk, Burak (14 March 2023). ["Google Cloud brings generative AI to developers, businesses, and governments"](#). [Archived](#) from the original on 17 November 2023. Retrieved 15 March 2023.
179. [^](#) Brynjolfsson, Erik; Li, Danielle; Raymond, Lindsey R. (April 2023), [Generative AI at Work](#) (Working Paper), Working Paper Series, [doi:10.3386/w31161](#), [archived](#) from the original on 28 March 2024, retrieved 21 January 2024
180. [^](#) ["Don't fear an AI-induced jobs apocalypse just yet"](#). *The Economist*. 6 March 2023. [Archived](#) from the original on 17 November 2023. Retrieved 14 March 2023.
181. [^](#) Coyle, Jake (27 September 2023). ["In Hollywood writers' battle against AI, humans win \(for now\)"](#). AP News. Associated Press. [Archived](#) from the original on 3 April 2024. Retrieved 26 January 2024.
182. [^](#) Harreis, H.; Koullias, T.; Roberts, Roger. ["Generative AI: Unlocking the future of fashion"](#). [Archived](#) from the original on 17 November 2023. Retrieved 14 March 2023.
183. [^](#) ["How Generative AI Can Augment Human Creativity"](#). *Harvard Business Review*. 16 June 2023. [ISSN 0017-8012](#). [Archived](#) from the original on 20 June 2023. Retrieved 20 June 2023.
184. [^](#) Hendrix, Justin (16 May 2023). ["Transcript: Senate Judiciary Subcommittee Hearing on Oversight of AI"](#). *techpolicy.press*. [Archived](#) from the original on 17 November 2023. Retrieved 19 May 2023.
185. [^](#) Simon, Felix M.; Altay, Sacha; Mercier, Hugo (18 October 2023). ["Misinformation reloaded? Fears about the impact of generative AI on misinformation are overblown"](#). *Harvard Kennedy School Misinformation Review*. [doi:10.37016/mr-2020-127](#). [S2CID 264113883](#). [Archived](#) from the original on 17 November 2023. Retrieved 16 November 2023.
186. [^](#) ["New AI systems collide with copyright law"](#). *BBC News*. 1 August 2023. Retrieved 28 September 2024.
187. [^](#) Poole, David; Mackworth, Alan (2023). [Artificial Intelligence, Foundations of Computational Agents](#) (3rd ed.). Cambridge University Press. [doi:10.1017/9781009258227](#). [ISBN 978-1-0092-5819-7](#). [Archived](#) from the original on 5 October 2024. Retrieved 5 October 2024.
188. [^](#) Russell, Stuart; Norvig, Peter (2020). [Artificial Intelligence: A Modern Approach](#) (4th ed.). Pearson. [ISBN 978-0-1346-1099-3](#).
189. [^](#) ["Why agents are the next frontier of generative AI"](#). *McKinsey Digital*. 24 July 2024. [Archived](#) from the original on 3 October 2024. Retrieved 10 August 2024.
190. [^](#) Figueiredo, Mayara Costa; Ankrah, Elizabeth; Powell, Jacquelyn E.; Epstein, Daniel A.; Chen, Yunan (12 January 2024). ["Powered by AI: Examining How AI](#)

- [Descriptions Influence Perceptions of Fertility Tracking Applications](#)". *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* **7** (4): 154:1–154:24. [doi:10.1145/3631414](#).
191. [^](#) Power, Jennifer; Pym, Tinonee; James, Alexandra; Waling, Andrea (5 July 2024). ["Smart Sex Toys: A Narrative Review of Recent Research on Cultural, Health and Safety Considerations"](#). *Current Sexual Health Reports*. **16** (3): 199–215. [doi:10.1007/s11930-024-00392-3](#). [ISSN 1548-3592](#).
 192. [^](#) Marcantonio, Tiffany L.; Avery, Gracie; Thrash, Anna; Leone, Ruschelle M. (10 September 2024). ["Large Language Models in an App: Conducting a Qualitative Synthetic Data Analysis of How Snapchat's "My AI" Responds to Questions About Sexual Consent, Sexual Refusals, Sexual Assault, and Sexting"](#). *The Journal of Sex Research*: 1–15. [doi:10.1080/00224499.2024.2396457](#). [ISSN 0022-4499](#). [PMID 39254628](#). [Archived](#) from the original on 9 December 2024. Retrieved 9 December 2024.
 193. [^](#) Hanson, Kenneth R.; Bolthouse, Hannah (2024). [""Replika Removing Erotic Role-Play Is Like Grand Theft Auto Removing Guns or Cars": Reddit Discourse on Artificial Intelligence Chatbots and Sexual Technologies"](#). *Socius: Sociological Research for a Dynamic World*. **10**. [doi:10.1177/23780231241259627](#). [ISSN 2378-0231](#).
 194. [^](#) Mania, Karolina (1 January 2024). ["Legal Protection of Revenge and Deepfake Porn Victims in the European Union: Findings From a Comparative Legal Study"](#). *Trauma, Violence, & Abuse*. **25** (1): 117–129. [doi:10.1177/15248380221143772](#). [ISSN 1524-8380](#). [PMID 36565267](#).
 195. [^](#) Singh, Suyesha; Nambiar, Vaishnavi (2024). ["Role of Artificial Intelligence in the Prevention of Online Child Sexual Abuse: A Systematic Review of Literature"](#). *Journal of Applied Security Research*. **19** (4): 586–627. [doi:10.1080/19361610.2024.2331885](#). [ISSN 1936-1610](#). [Archived](#) from the original on 9 December 2024. Retrieved 9 December 2024.
 196. [^](#) Razi, Afsaneh; Kim, Seunghyun; Alsoubai, Ashwaq; Stringhini, Gianluca; Solorio, Thamar; De Choudhury, Munmun; Wisniewski, Pamela J. (13 October 2021). ["A Human-Centered Systematic Literature Review of the Computational Approaches for Online Sexual Risk Detection"](#). *Proceedings of the ACM on Human-Computer Interaction*. **5** (CSCW2): 1–38. [doi:10.1145/3479609](#). [ISSN 2573-0142](#). [Archived](#) from the original on 9 December 2024. Retrieved 9 December 2024.
 197. [^](#) Ransbotham, Sam; Kiron, David; Gerbert, Philipp; Reeves, Martin (6 September 2017). ["Reshaping Business With Artificial Intelligence"](#). *MIT Sloan Management Review*. [Archived](#) from the original on 13 February 2024.
 198. [^](#) Sun, Yuran; Zhao, Xilei; Lovreglio, Ruggiero; Kuligowski, Erica (1 January 2024), Naser, M. Z. (ed.), ["8 – AI for large-scale evacuation modeling: promises and challenges"](#), *Interpretable Machine Learning for the Analysis, Design, Assessment, and Informed Decision Making for Civil Infrastructure*, Woodhead Publishing Series in Civil and Structural Engineering, Woodhead Publishing, pp. 185–204, [ISBN 978-0-1282-4073-1](#), [archived](#) from the original on 19 May 2024, retrieved 28 June 2024.

199. [^ Gomaa, Islam; Adelzadeh, Masoud; Gwynne, Steven; Spencer, Bruce; Ko, Yoon; Bénichou, Nouredine; Ma, Chunyun; Elsagan, Nour; Duong, Dana; Zalog, Ehab; Kinatader, Max \(1 November 2021\). "A Framework for Intelligent Fire Detection and Evacuation System". *Fire Technology*. **57** \(6\): 3179–3185. doi:10.1007/s10694-021-01157-3. ISSN 1572-8099. Archived from the original on 5 October 2024. Retrieved 5 October 2024.](#)
200. [^ Zhao, Xilei; Lovreglio, Ruggiero; Nilsson, Daniel \(1 May 2020\). "Modelling and interpreting pre-evacuation decision-making using machine learning". *Automation in Construction*. **113**: 103140. doi:10.1016/j.autcon.2020.103140. hdl:10179/17315. ISSN 0926-5805. Archived from the original on 19 May 2024. Retrieved 5 October 2024.](#)
201. [^ "India's latest election embraced AI technology. Here are some ways it was used constructively". *PBS News*. 12 June 2024. Archived from the original on 17 September 2024. Retrieved 28 October 2024.](#)
202. [^ Müller, Vincent C. \(30 April 2020\). "Ethics of Artificial Intelligence and Robotics". *Stanford Encyclopedia of Philosophy Archive*. Archived from the original on 5 October 2024. Retrieved 5 October 2024.](#)
203. [^ Simonite \(2016\).](#)
204. [^ Russell & Norvig \(2021\)](#), p. 987.
205. [^ Laskowski \(2023\).](#)
206. [^ GAO \(2022\).](#)
207. [^ Valinsky \(2019\).](#)
208. [^ Russell & Norvig \(2021\)](#), p. 991.
209. [^ Russell & Norvig \(2021\)](#), pp. 991–992.
210. [^ Christian \(2020\)](#), p. 63.
211. [^ Vincent \(2022\).](#)
212. [^ Kopel, Matthew. "Copyright Services: Fair Use". *Cornell University Library*. Archived from the original on 26 September 2024. Retrieved 26 April 2024.](#)
213. [^ Burgess, Matt. "How to Stop Your Data From Being Used to Train AI". *Wired*. ISSN 1059-1028. Archived from the original on 3 October 2024. Retrieved 26 April 2024.](#)
214. [^ Reisner \(2023\).](#)
215. [^ Alter & Harris \(2023\).](#)
216. [^ "Getting the Innovation Ecosystem Ready for AI. An IP policy toolkit" \(PDF\). *WIPO*.](#)
217. [^ Hammond, George \(27 December 2023\). "Big Tech is spending more than VC firms on AI startups". *Ars Technica*. Archived from the original on 10 January 2024.](#)

218. [^] Wong, Matteo (24 October 2023). ["The Future of AI Is GOMA"](#). The Atlantic. [Archived](#) from the original on 5 January 2024.
219. [^] ["Big tech and the pursuit of AI dominance"](#). The Economist. 26 March 2023. [Archived](#) from the original on 29 December 2023.
220. [^] Fung, Brian (19 December 2023). ["Where the battle to dominate AI may be won"](#). CNN Business. [Archived](#) from the original on 13 January 2024.
221. [^] Metz, Cade (5 July 2023). ["In the Age of A.I., Tech's Little Guys Need Big Friends"](#). The New York Times. [Archived](#) from the original on 8 July 2024. Retrieved 5 October 2024.
222. [^] ["Electricity 2024 – Analysis"](#). IEA. 24 January 2024. Retrieved 13 July 2024.
223. [^] Calvert, Brian (28 March 2024). ["AI already uses as much energy as a small country. It's only the beginning"](#). Vox. New York, New York. [Archived](#) from the original on 3 July 2024. Retrieved 5 October 2024.
224. [^] Halper, Evan; O'Donovan, Caroline (21 June 2024). ["AI is exhausting the power grid. Tech firms are seeking a miracle solution"](#). Washington Post.
225. [^] Davenport, Carly. ["AI Data Centers and the Coming YS Power Demand Surge"](#) (PDF). Goldman Sachs. Archived from [the original](#) (PDF) on 26 July 2024. Retrieved 5 October 2024.
226. [^] Ryan, Carol (12 April 2024). ["Energy-Guzzling AI Is Also the Future of Energy Savings"](#). Wall Street Journal. Dow Jones.
227. [^] Hiller, Jennifer (1 July 2024). ["Tech Industry Wants to Lock Up Nuclear Power for AI"](#). Wall Street Journal. Dow Jones. [Archived](#) from the original on 5 October 2024. Retrieved 5 October 2024.
228. [^] Kendall, Tyler (28 September 2024). ["Nvidia's Huang Says Nuclear Power an Option to Feed Data Centers"](#). Bloomberg.
229. [^] Halper, Evan (20 September 2024). ["Microsoft deal would reopen Three Mile Island nuclear plant to power AI"](#). Washington Post.
230. [^] Hiller, Jennifer (20 September 2024). ["Three Mile Island's Nuclear Plant to Reopen, Help Power Microsoft's AI Centers"](#). Wall Street Journal. Dow Jones. [Archived](#) from the original on 5 October 2024. Retrieved 5 October 2024.
231. [^] [Jump up to:](#)^{a b c} Niva Yadav (19 August 2024). ["Taiwan to stop large data centers in the North, cites insufficient power"](#). DatacenterDynamics. [Archived](#) from the original on 8 November 2024. Retrieved 7 November 2024.
232. [^] [Jump up to:](#)^{a b} Mochizuki, Takashi; Oda, Shoko (18 October 2024). ["エヌビディア出資の日本企業、原発近くでAIデータセンター新設検討"](#). Bloomberg (in Japanese). [Archived](#) from the original on 8 November 2024. Retrieved 7 November 2024.

233. ^a ^b [Jump up to:](#) Naureen S Malik and Will Wade (5 November 2024). ["Nuclear-Hungry AI Campuses Need New Plan to Find Power Fast"](#). Bloomberg.
234. ^a [Nicas \(2018\)](#).
235. ^a Rainie, Lee; Keeter, Scott; Perrin, Andrew (22 July 2019). ["Trust and Distrust in America"](#). Pew Research Center. [Archived](#) from the original on 22 February 2024.
236. ^a [Williams \(2023\)](#).