Artificial intelligence

"AI" redirects here. For other uses, see AI (disambiguation), Artificial intelligence (disambiguation), and Intelligent agent.

Part of a series on
Artificial intelligence

Major goals
Artificial general intelligence · Planning · Computer vision · General game playing · Knowledge reasoning · Machine learning · Natural language processing · Robotics · AI safety

Approaches
Symbolic · Deep learning · Bayesian networks · Evolutionary algorithms · Situated approach · Hybrid intelligent systems · Systems integration

Philosophy
Chinese room · Friendly AI · Control problem/Takeover · Ethics · Existential risk · Turing test

History
Timeline · Progress · AI winter

Technology
Applications · Projects · Programming languages

Glossary
Glossary

v · t · e

Artificial intelligence (**AI**) is intelligence—perceiving, synthesizing, and inferring information—demonstrated by machines, as opposed to intelligence displayed by non-human animals or by humans. Example tasks in which this is done include speech recognition, computer vision, translation between (natural) languages, as well as other mappings of inputs.

AI applications include advanced web search engines (e.g., Google Search), recommendation systems (used by YouTube, Amazon, and Netflix), understanding human speech (such as Siri and Alexa), self-driving cars (e.g., Waymo), generative or creative tools (ChatGPT and AI art), automated decision-making, and competing at the highest level in strategic game systems (such as chess and Go).[1]

As machines become increasingly capable, tasks considered to require "intelligence" are often removed from the definition of AI, a phenomenon known as the AI effect.[2] For instance, optical character recognition is frequently excluded from things considered to be AI, having become a routine technology.[3]

Artificial intelligence was founded as an academic discipline in 1956, and in the years since it has experienced several waves of optimism,[4][5] followed by disappointment and the loss of funding (known as an "AI winter"),[6][7] followed by new approaches, success, and renewed funding.[5][8] AI research has tried and discarded many different approaches, including simulating the brain, modeling human problem solving, formal logic, large databases of knowledge, and imitating animal behavior. In the first decades of the 21st century, highly mathematical and statistical machine learning has dominated the field, and this technique has proved highly successful, helping to solve many challenging problems throughout industry and academia.[8][9]

The various sub-fields 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 the ability to move and manipulate objects.[a] General intelligence (the ability to solve an arbitrary problem) is among the field's long-term goals.[10] To solve these problems, AI researchers have adapted and integrated a wide range of problem-solving techniques, including search and mathematical optimization, formal logic, artificial neural networks, and methods based on statistics, probability, and economics. AI also draws upon computer science, psychology, linguistics, philosophy, and many other fields.

The field was founded on the assumption that human intelligence "can be so precisely described that a machine can be made to simulate it".[b] This raised philosophical arguments about the mind and the ethical consequences of creating artificial beings endowed with human-like intelligence; these issues have previously been explored by myth, fiction, and philosophy since antiquity.[12] Computer scientists and philosophers have since suggested that AI may become an existential risk to humanity if its rational capacities are not steered towards beneficial goals.[c] The term artificial intelligence has also been criticized for overhyping AI's true technological capabilities.[13][14][15]

History

Main articles: History of artificial intelligence and Timeline of artificial intelligence

Silver didrachma from Crete depicting Talos, an ancient mythical automaton with artificial intelligence

Artificial beings with intelligence appeared as storytelling devices in antiquity,[16] and have been common in fiction, as in Mary Shelley's *Frankenstein* or Karel Čapek's *R.U.R.*[17] These characters and their fates raised many of the same issues now discussed in the ethics of artificial intelligence.[18]

The study of mechanical or "formal" reasoning began with philosophers and mathematicians in antiquity. The study of mathematical 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 act of mathematical deduction. This insight that digital computers can simulate any process of formal reasoning is known as the Church–Turing thesis.[19] This, along with concurrent discoveries in neurobiology, information theory and cybernetics, led researchers to consider the possibility of building an electronic brain.[20] The first work that is now generally recognized as AI was McCullouch and Pitts' 1943 formal design for Turing-complete "artificial neurons".[21]

By the 1950s, two visions for how to achieve machine intelligence emerged. One vision, known as Symbolic AI or GOFAI, was to use computers to create a symbolic representation of the world and systems that could reason about the world. Proponents included Allen Newell, Herbert A. Simon, and Marvin Minsky. Closely associated with this approach was the "heuristic search" approach, which likened intelligence to a problem of exploring a space of possibilities for answers.

The second vision, known as the connectionist approach, sought to achieve intelligence through learning. Proponents of this approach, most prominently Frank Rosenblatt, sought to connect Perceptron in ways inspired by connections of neurons.[22] James Manyika and others have compared the two approaches to the mind (Symbolic AI) and the brain (connectionist). Manyika argues that symbolic approaches dominated the push for artificial intelligence in this period, due in part to its connection to intellectual traditions of Descartes, Boole, Gottlob Frege, Bertrand Russell, and others. Connectionist approaches based on cybernetics or artificial neural networks were pushed to the background but have gained new prominence in recent decades.[23]

The field of AI research was born at a workshop at Dartmouth College in 1956.[d][26] The attendees became the founders and leaders of AI research.[e] They and their students produced programs that the press described as "astonishing":[f] computers were learning checkers strategies, solving word problems in algebra, proving logical theorems and speaking English.[g][28]

By the middle of the 1960s, research in the U.S. was heavily funded by the Department of Defense[29] and laboratories had been established around the world.[30]

Researchers in the 1960s and the 1970s were convinced that symbolic approaches would eventually succeed in creating a machine with artificial general intelligence and considered this the goal of their field.[31] Herbert Simon predicted, "machines will be capable, within twenty years, of doing any work a man can do".[32] Marvin Minsky agreed, writing, "within a generation ... the problem of creating 'artificial intelligence' will substantially be solved".[33]

They had failed to recognize the difficulty of some of the remaining tasks. Progress slowed and in 1974, in response to the criticism of Sir James Lighthill[34] and ongoing pressure from the US Congress to fund more productive projects, both the U.S. and British governments cut off exploratory research in AI. The next few years would later be called an "AI winter", a period when obtaining funding for AI projects was difficult.[6]

In the early 1980s, AI research was revived by the commercial success of expert systems,[35] 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.[5] 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.[7]

Many researchers began to doubt that the symbolic approach would be able to imitate all the processes of human cognition, especially perception, robotics, learning and pattern recognition. A number of researchers began to look into "sub-symbolic" approaches to specific AI problems.[36] Robotics researchers, such as Rodney Brooks, rejected symbolic AI and focused on the basic engineering problems that would allow robots to move, survive, and learn their environment.[h]

Interest in neural networks and "connectionism" was revived by Geoffrey Hinton, David Rumelhart and others in the middle of the 1980s.[41] Soft computing tools were developed in the 1980s, such as neural networks, fuzzy systems, Grey system theory, evolutionary computation and many tools drawn from statistics or mathematical optimization.

AI gradually restored its reputation in the late 1990s and early 21st century by finding specific solutions to specific problems. The narrow focus allowed researchers to produce verifiable results, exploit more mathematical methods, and collaborate with other fields (such as statistics, economics and mathematics).[42] By 2000, solutions developed by AI researchers were being widely used, although in the 1990s they were rarely described as "artificial intelligence".[9]

Faster computers, algorithmic improvements, and access to large amounts of data enabled advances in machine learning and perception; data-hungry deep learning methods started to dominate accuracy benchmarks around 2012.[43] According to Bloomberg's Jack Clark, 2015 was a landmark year for artificial intelligence, with the number of software projects that use AI within Google increased from a "sporadic usage" in 2012 to more than 2,700 projects.[i] He attributed this to an increase in affordable neural networks, due to a rise in cloud computing infrastructure and to an increase in research tools and datasets.[8]

In a 2017 survey, one in five companies reported they had "incorporated AI in some offerings or processes".[44] The amount of research into AI (measured by total publications) increased by 50% in the years 2015–2019.[45]

Numerous academic researchers became concerned that AI was no longer pursuing the original goal of creating versatile, fully intelligent machines. Much of current research involves statistical AI, which is overwhelmingly used to solve specific problems, even highly successful techniques such as deep learning. This concern has led to the subfield of artificial general intelligence (or "AGI"), which had several well-funded institutions by the 2010s.[10]

In April 2023, computer scientist Jaron Lanier published an alternative view of AI in *The New Yorker* as less intelligent than the name, and popular culture, may suggest. Lanier concludes his essay as follows: "Think of people. People are the answer to the problems of bits."[46][47]

Goals

The general problem of simulating (or creating) intelligence has been broken down into sub-problems. These consist of particular traits or capabilities that researchers expect an intelligent system to display. The traits described below have received the most attention.[a]

Reasoning, problem-solving

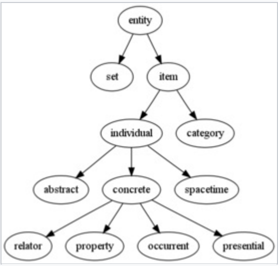
Early researchers developed algorithms that imitated step-by-step reasoning that humans use when they solve puzzles or make logical deductions.[48] By the late 1980s and 1990s, AI research had developed methods for dealing with uncertain or incomplete information, employing concepts from probability and economics.[49]

Many of these algorithms proved to be insufficient for solving large reasoning problems because they experienced a "combinatorial explosion": they became exponentially slower as the problems grew larger.[50] 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.[51]

Knowledge representation

Main articles: Knowledge representation, Commonsense knowledge, Description logic, and Ontology

Knowledge representation and knowledge engineering[52] allow AI programs to answer questions intelligently and make deductions about real-world facts.



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

A representation of "what exists" is an ontology: the set of objects, relations, concepts, and properties formally described so that software agents can interpret them.[53] The most general ontologies are called upper ontologies, which attempt to provide a foundation for all other knowledge and act as mediators between domain ontologies that cover specific knowledge about a particular knowledge domain (field of interest or area of concern). A truly intelligent program would also need access to commonsense knowledge; the set of facts that an average person knows. The semantics of an ontology is typically represented in description logic, such as the Web Ontology Language.[54]

AI research has developed tools to represent specific domains, such as objects, properties, categories and relations between objects;[54] situations, events, states and time;[55] causes and effects;[56] knowledge about knowledge (what we know about what other people know);.[57] default reasoning (things that humans assume are true until they are told differently and will remain true even when other facts are changing);[58] as well as other domains. Among the most difficult problems in AI are: the breadth of commonsense knowledge (the number of atomic facts that the average person knows is enormous);[59] 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).[51]

Formal knowledge representations are used in content-based indexing and retrieval,[60] scene interpretation,[61] clinical decision support,[62] knowledge discovery (mining "interesting" and actionable inferences from large databases),[63] and other areas.[64]

Learning

Main article: Machine learning

Machine learning (ML), a fundamental concept of AI research since the field's inception,[j] is the study of computer algorithms that improve automatically through experience.[k]

Unsupervised learning finds patterns in a stream of input.

Supervised learning requires a human to label the input data first, and comes in two main varieties: classification and numerical regression. Classification is used to determine what category something belongs in – the program sees a number of examples of things from several categories and will learn to classify new inputs. Regression is the attempt to produce a function that describes the relationship between inputs and outputs and predicts how the outputs should change as the inputs change. Both classifiers and regression learners can be viewed as "function approximators" trying to learn an unknown (possibly implicit) function; for example, a spam classifier can be viewed as learning a function that maps from the text of an email to one of two categories, "spam" or "not spam".[68]

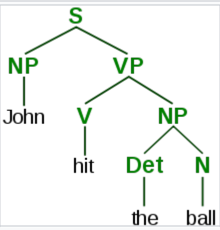
In reinforcement learning the agent is rewarded for good responses and punished for bad ones. The agent classifies its responses to form a strategy for operating in its problem space.[69]

Transfer learning is when the knowledge gained from one problem is applied to a new problem.[70]

Computational learning theory can assess learners by computational complexity, by sample complexity (how much data is required), or by other notions of optimization.[71]

Natural language processing

Main article: Natural language processing



A parse tree represents the syntactic structure of a sentence according to some formal grammar.

Natural language processing (NLP)[72] allows machines to read and understand human language. A sufficiently powerful natural language processing system would enable natural-language user interfaces and the acquisition of knowledge directly from human-written sources, such as newswire texts. Some straightforward applications of NLP include information retrieval, question answering and machine translation.[73]

Symbolic AI used formal syntax to translate the deep structure of sentences into logic. This failed to produce useful applications, due to the intractability of logic[50] and the breadth of commonsense knowledge.[59] Modern statistical techniques include co-occurrence frequencies (how often one word appears near another), "Keyword spotting" (searching for a particular word to retrieve information), transformer-based deep learning (which finds patterns in text), and others.[74] They have achieved acceptable accuracy at the page or paragraph level, and, by 2019, could generate coherent text.[75]

Perception

Main articles: Machine perception, Computer vision, and Speech recognition

Machine perception[76] is the ability to use input from sensors (such as cameras, microphones, wireless signals, and active lidar, sonar, radar, and tactile sensors) to deduce aspects of the world. Applications include speech recognition,[77] facial recognition, and object recognition.[78] Computer vision is the ability to analyze visual input.[79]


Feature detection (pictured: edge detection) helps AI compose informative abstract structures out of raw data.

Social intelligence

Main article: Affective computing

Affective computing is an interdisciplinary umbrella that comprises systems that recognize, interpret, process or simulate human feeling, emotion and mood.[81] 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 how intelligent existing computer agents actually are.[82] Moderate successes related to affective computing include textual sentiment analysis and, more recently, multimodal sentiment analysis), wherein AI classifies the affects displayed by a videotaped subject.[83]


Kismet, a robot with rudimentary social skills[80]

General intelligence

Main article: Artificial general intelligence

A machine with general intelligence can solve a wide variety of problems with breadth and versatility similar to human intelligence. There are several competing ideas about how to develop artificial general intelligence. Hans Moravec and Marvin Minsky argue that work in different individual domains can be incorporated into an advanced multi-agent system or cognitive architecture with general intelligence.[84] Pedro Domingos hopes that there is a conceptually straightforward, but mathematically difficult, "master algorithm" that could lead to AGI.[85] Others believe that anthropomorphic features like an artificial brain[86] or simulated child development[l] will someday reach a critical point where general intelligence emerges.

Tools

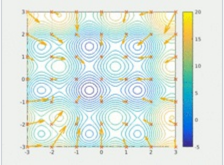
Search and optimization

Main articles: Search algorithm, Mathematical optimization, and Evolutionary computation

AI can solve many problems by intelligently searching through many possible solutions.[87] Reasoning can be reduced to performing a search. For example, logical proof can be viewed as searching for a path that leads from premises to conclusions, where each step is the application of an inference rule.[88] Planning algorithms search through trees of goals and subgoals, attempting to find a path to a target goal, a process called means-ends analysis.[89] Robotics algorithms for moving limbs and grasping objects use local searches in configuration space.[90]

Simple exhaustive searches[91] 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. The solution, for many problems, is to use "heuristics" or "rules of thumb" that prioritize choices in favor of those more likely to reach a goal and to do so in a shorter number of steps. In some search methodologies, heuristics can also serve to eliminate some choices unlikely to lead to a goal (called "pruning the search tree"). Heuristics supply the program with a "best guess" for the path on which the solution lies.[92] Heuristics limit the search for solutions into a smaller sample size.[93]

A very different kind of search came to prominence in the 1990s, based on the mathematical theory of optimization. For many problems, it is possible to begin the search with some form of a guess and then refine the guess incrementally until no more refinements can be made. These algorithms can be visualized as blind hill climbing: we begin the search at a random point on the landscape, and then, by jumps or steps, we keep moving our guess uphill, until we reach the top. Other related optimization algorithms include random optimization, beam search and metaheuristics like simulated annealing.[94] Evolutionary computation uses a form of optimization search. For example, they may begin with a population of organisms (the guesses) and then allow them to mutate and recombine, selecting only the fittest to survive each generation (refining the guesses). Classic evolutionary algorithms include genetic algorithms, gene expression programming, and genetic programming.[95] Alternatively, 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).[96]



A particle swarm seeking the global minimum

Logic

Main articles: Logic programming and Automated reasoning

Logic[97] is used for knowledge representation and problem-solving, but it can be applied to other problems as well. For example, the satplan algorithm uses logic for planning[98] and inductive logic programming is a method for learning.[99]

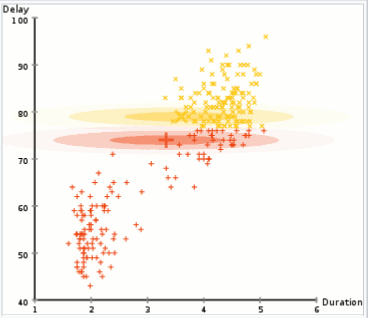
Several different forms of logic are used in AI research. Propositional logic[100] involves truth functions such as "or" and "not". First-order logic[101] adds quantifiers and predicates and can express facts about objects, their properties, and their relations with each other. Fuzzy logic assigns a "degree of truth" (between 0 and 1) to vague statements such as "Alice is old" (or rich, or tall, or hungry), that are too linguistically imprecise to be completely true or false.[102] Default logics, non-monotonic logics and circumscription are forms of logic designed to help with default reasoning and the qualification problem.[58] Several extensions of logic have been designed to handle specific domains of knowledge, such as description logics;[54] situation calculus, event calculus and fluent calculus (for representing events and time);[55] causal calculus;[56] belief calculus (belief revision); and modal logics.[57] Logics to model contradictory or inconsistent statements arising in multi-agent systems have also been designed, such as paraconsistent logics.[103]

Probabilistic methods for uncertain reasoning

Main articles: Bayesian network, Hidden Markov model, Kalman filter, Particle filter, Decision theory, and Utility theory

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.[104] Bayesian networks[105] are a very general tool that can be used for various problems, including reasoning (using the Bayesian inference algorithm),[m][107] learning (using the expectation-maximization algorithm),[n][109] planning (using decision networks)[110] and perception (using dynamic Bayesian networks).[111] Probabilistic algorithms can also be used for filtering, prediction, smoothing and finding explanations for streams of data, helping perception systems to analyze processes that occur over time (e.g., hidden Markov models or Kalman filters).[111]

A key concept from the science of economics is "utility", a measure of how valuable something is to an intelligent agent. Precise mathematical tools have been developed that analyze how an agent can make choices and plan, using decision theory, decision analysis,[112] and information value theory.[113] These tools include models such as Markov decision processes,[114] dynamic decision networks,[111] game theory and mechanism design.[115]



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

Main articles: Statistical classification and Machine learning

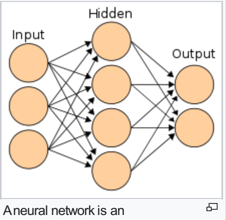
The simplest AI applications can be divided into two types: classifiers ("if shiny then diamond") and controllers ("if diamond then pick up"). Controllers do, however, also classify conditions before inferring actions, and therefore classification forms a central part of many AI systems. Classifiers are functions that use pattern matching to determine the closest match. They can be tuned according to examples, making them very attractive for use in AI. These examples are known as observations or patterns. In supervised learning, each pattern belongs to a certain predefined class. A class is a decision that has to be made. 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.[116]

A classifier can be trained in various ways; there are many statistical and machine learning approaches. The decision tree is the simplest and most widely used symbolic machine learning algorithm.[117] K-nearest neighbor algorithm was the most widely used analogical AI until the mid-1990s.[118] Kernel methods such as the support vector machine (SVM) displaced k-nearest neighbor in the 1990s.[119] The naive Bayes classifier is reportedly the "most widely used learner"[120] at Google, due in part to its scalability.[121] Neural networks are also used for classification.[122]

Classifier performance depends greatly on the characteristics of the data to be classified, such as the dataset size, distribution of samples across classes, dimensionality, and the level of noise. Model-based classifiers perform well if the assumed model is an extremely good fit for the actual data. Otherwise, if no matching model is available, and if accuracy (rather than speed or scalability) is the sole concern, conventional wisdom is that discriminative classifiers (especially SVM) tend to be more accurate than model-based classifiers such as "naive Bayes" on most practical data sets.[123]

Artificial neural networks

Main articles: Artificial neural network and Connectionism



A neural network is an interconnected group of nodes, akin to the vast network of neurons in the human brain.

Neural networks[122] were inspired by the architecture of neurons in the human brain. A simple "neuron" *N* accepts input from other neurons, each of which, when activated (or "fired"), casts a weighted "vote" for or against whether neuron *N* should itself activate. Learning requires an algorithm to adjust these weights based on the training data; one simple algorithm (dubbed "fire together, wire together") is to increase the weight between two connected neurons when the activation of one triggers the successful activation of another. Neurons have a continuous spectrum of activation; in addition, neurons can process inputs in a nonlinear way rather than weighing straightforward votes.

Modern neural networks model complex relationships between inputs and outputs and find patterns in data. They can learn continuous functions and even digital logical operations. Neural networks can be viewed as a type of mathematical optimization – they perform gradient descent on a multi-dimensional topology that was created by training the network. The most common training technique is the backpropagation algorithm.[124] Other learning techniques for neural networks are Hebbian learning ("fire together, wire together"), GMDH or competitive learning.[125]

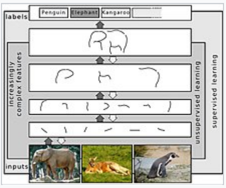
The main categories of networks are acyclic or feedforward neural networks (where the signal passes in only one direction) and recurrent neural networks (which allow feedback and short-term memories of previous input events). Among the most popular feedforward networks are perceptrons, multi-layer perceptrons and radial basis networks.[126]

Deep learning

Deep learning[128] 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 or letters or faces.[129] Deep learning has drastically improved the performance of programs in many important subfields of artificial intelligence, including computer vision, speech recognition, image classification[130] and others.

Deep learning often uses convolutional neural networks for many or all of its layers. In a convolutional layer, each neuron receives input from only a restricted area of the previous layer called the neuron's receptive field. This can substantially reduce the number of weighted connections between neurons,[131] and creates a hierarchy similar to the organization of the animal visual cortex.[132]

In a recurrent neural network (RNN) the signal will propagate through a layer more than once;[133] thus, an RNN is an example of deep learning.[134] RNNs can be trained by gradient descent,[135] however long-term gradients which are back-propagated can "vanish" (that is, they can tend to zero) or "explode" (that is, they can tend to infinity), known as the vanishing gradient problem.[136] The long short term memory (LSTM) technique can prevent this in most cases.[137]



Representing images on multiple layers of abstraction in deep learning[127]

Specialized languages and hardware

Main articles: Programming languages for artificial intelligence and Hardware for artificial intelligence

Specialized languages for artificial intelligence have been developed, such as Lisp, Prolog, TensorFlow and many others. Hardware developed for AI includes AI accelerators and neuromorphic computing.

Applications

Main article: Applications of artificial intelligence
See also: Embodied cognition and Legal informatics

AI is relevant to any intellectual task.[138] Modern artificial intelligence techniques are pervasive and are too numerous to list here.[139] Frequently, when a technique reaches mainstream use, it is no longer considered artificial intelligence; this phenomenon is described as the AI effect.[140]

For this project of the artist Joseph Ayerle the AI had to learn the typical patterns in the colors and brushstrokes of Renaissance painter Raphael. The portrait shows the face of the actress Ornella Muti, "painted" by AI in the style of Raphael.

In the 2010s, AI applications were at the heart of the most commercially successful areas of computing, and have become a ubiquitous feature of daily life. AI is used in search engines (such as Google Search), targeting online advertisements,[141] recommendation systems (offered by Netflix, YouTube or Amazon), driving internet traffic,[142][143] targeted advertising (AdSense, Facebook), virtual assistants (such as Siri or Alexa),[144] 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), image labeling (used by Facebook, Apple's iPhoto and TikTok) , spam filtering and chatbots (such as Chat GPT).

There are also thousands of successful AI applications used to solve problems for specific industries or institutions. A few examples are energy storage,[145] deepfakes,[146] medical diagnosis, military logistics, foreign policy,[147] or supply chain management.

Game playing has been a test of AI's strength since the 1950s. Deep Blue became the first computer chess-playing system to beat a reigning world chess champion, Garry Kasparov, on 11 May 1997.[148] 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.[149] 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.[150] Other programs handle imperfect-information games; such as for poker at a superhuman level, Pluribus[o] and Cepheus.[152] DeepMind in the 2010s developed a "generalized artificial intelligence" that could learn many diverse Atari games on its own.[153]

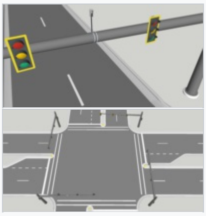
By 2020, Natural Language Processing systems such as the enormous GPT-3 (then by far the largest artificial neural network) were matching human performance on pre-existing benchmarks, albeit without the system attaining a commonsense understanding of the contents of the benchmarks.[154] DeepMind's AlphaFold 2 (2020) demonstrated the ability to approximate, in hours rather than months, the 3D structure of a protein.[155] Other applications predict the result of judicial decisions,[156] create art (such as poetry or painting) and prove mathematical theorems.

AI content detector tools are software applications that use artificial intelligence (AI) algorithms to analyze and detect specific types of content in digital media, such as text, images, and videos. These tools are commonly used to identify inappropriate content, such as speech errors, violent or sexual images, and spam, among others.

Some benefits of using AI content detector tools[157] include improved efficiency and accuracy in detecting inappropriate content, increased safety and security for users, and reduced legal and reputational risks for websites and platforms.

Smart traffic lights

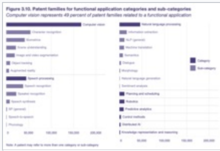
Smart traffic lights have been developed at Carnegie Mellon since 2009. Professor Stephen Smith has started a company since then Surtrac that has installed smart traffic control systems in 22 cities. It costs about $20,000 per intersection to install. Drive time has been reduced by 25% and traffic jam waiting time has been reduced by 40% at the intersections it has been installed.[158]



Artificially intelligent traffic lights use cameras with radar, ultrasonic acoustic location sensors, and predictive algorithms to improve traffic flow

Intellectual property

In 2019, WIPO reported that AI was the most prolific emerging technology in terms of the number of patent applications and granted patents, the Internet of things was estimated to be the largest in terms of market size. It was followed, again in market size, by big data technologies, robotics, AI, 3D printing and the fifth generation of mobile services (5G).[159] Since AI emerged in the 1950s, 340,000 AI-related patent applications were filed by innovators and 1.6 million scientific papers have been published by researchers, with the majority of all AI-related patent filings published since 2013. Companies represent 26 out of the top 30 AI patent applicants, with universities or public research organizations accounting for the remaining four.[160] The ratio of scientific papers to inventions has significantly decreased from 8:1 in 2010 to 3:1 in 2016, which is attributed to be indicative of a shift from theoretical research to the use of AI technologies in commercial products and services. Machine learning is the dominant AI technique disclosed in patents and is included in more than one-third of all identified inventions (134,777 machine learning patents filed for a total of 167,038 AI patents filed in 2016), with computer vision being the most popular functional application. AI-related patents not only disclose AI techniques and applications, they often also refer to an application field or industry. Twenty application fields were identified in 2016 and included, in order of magnitude: telecommunications (15 percent), transportation (15 percent), life and medical sciences (12 percent), and personal devices, computing and human–computer interaction (11 percent). Other sectors included banking, entertainment, security, industry and manufacturing, agriculture, and networks (including social networks, smart cities and the Internet of things). IBM has the largest portfolio of AI patents with 8,290 patent applications, followed by Microsoft with 5,930 patent applications.[160]



AI patent families for functional application categories and sub categories. Computer vision represents 49 percent of patent families related to a functional application in 2016.

Philosophy

Main article: Philosophy of artificial intelligence

Defining artificial intelligence

Main articles: Turing test, Intelligent agent, Dartmouth workshop, and Synthetic intelligence

Alan Turing wrote in 1950 "I propose to consider the question 'can machines think'?"[161] He advised changing the question from whether a machine "thinks", to "whether or not it is possible for machinery to show intelligent behaviour".[161] He devised the Turing test, which measures the ability of a machine to simulate human conversation.[162] 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[p] but "it is usual to have a polite convention that everyone thinks"[163]

Russell and Norvig agree with Turing that AI must be defined in terms of "acting" and not "thinking".[164] However, they are critical that the test compares machines to *people*. "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.' "[165] AI founder John McCarthy agreed, writing that "Artificial intelligence is not, by definition, simulation of human intelligence".[166]

McCarthy defines intelligence as "the computational part of the ability to achieve goals in the world."[167] Another AI founder, Marvin Minsky similarly defines it as "the ability to solve hard problems".[168] 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.

A definition that has also been adopted by Google[169][*better source needed*] - major practitionary in the field of AI. This definition stipulated the ability of systems to synthesize information as the manifestation of intelligence, similar to the way it is defined in biological intelligence.

Evaluating approaches to AI

No established unifying theory or paradigm has guided AI research for most of its history.[q] 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, neat, soft and narrow (see below). Critics argue that these questions may have to be revisited by future generations of AI researchers.

Symbolic AI and its limits

Main articles: Symbolic AI, Physical symbol systems hypothesis, Moravec's paradox, and Hubert Dreyfus's views on artificial intelligence

Symbolic AI (or "GOFAI")[171] 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."[172]

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.[173] 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.[174] Although his arguments had been ridiculed and ignored when they were first presented, eventually, AI research came to agree.[r][51]

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,[176][177] 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 (especially in areas like common sense reasoning). This issue was actively discussed in the 70s and 80s,[178] but in the 1990s mathematical methods and solid scientific standards became the norm, a transition that Russell and Norvig termed "the victory of the neats".[179]

Soft vs. hard computing

Main article: Soft computing

Finding a provably correct or optimal solution is intractable for many important problems.[50] 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 80s and most successful AI programs in the 21st century are examples of soft computing with neural networks.

Narrow vs. general AI

Main article: Artificial general intelligence

AI researchers are divided as to whether to pursue the goals of artificial general intelligence and superintelligence (general AI) 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.[180][181] 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 experimental 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. Stuart Russell and Peter Norvig observe that most AI researchers "don't care about the [philosophy of AI] – as long as the program works, they don't care whether you call it a simulation of intelligence or real intelligence."[182] 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.[183] 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. Human information processing is easy to explain, however, 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*.[184]

Computationalism and functionalism

Main articles: Computational theory of mind, Functionalism (philosophy of mind), and Chinese room

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.[185]

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."[s] Searle counters this assertion with his Chinese room argument, which attempts to show that, even if a machine perfectly simulates human behavior, there is still no reason to suppose it also has a mind.[188]

Robot rights

Main article: Robot rights

If a machine has a mind and subjective experience, then it may also have sentience (the ability to feel), and if so, then it could also *suffer*, and thus it would be entitled to certain rights.[189] Any hypothetical robot rights would lie on a spectrum with animal rights and human rights.[190] This issue has been considered in fiction for centuries,[191] and is now being considered by, for example, California's Institute for the Future; however, critics argue that the discussion is premature.[192]

Future

Superintelligence

Main articles: Superintelligence, Technological singularity, and Transhumanism

A superintelligence, hyperintelligence, or superhuman intelligence, is a hypothetical agent that would possess intelligence far surpassing that of the brightest and most gifted human mind. *Superintelligence* may also refer to the form or degree of intelligence possessed by such an agent.[181]

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 recursive self-improvement.[193] Its intelligence would increase exponentially in an intelligence explosion and could dramatically surpass humans. Science fiction writer Vernor Vinge named this scenario the "singularity".[194] Because it is difficult or impossible to know the limits of intelligence or the capabilities of superintelligent machines, the technological singularity is an occurrence beyond which events are unpredictable or even unfathomable.[195]

Robot designer Hans Moravec, cyberneticist Kevin Warwick, and inventor Ray Kurzweil have predicted that humans and machines will merge in the future into cyborgs that are more capable and powerful than either. This idea, called transhumanism, has roots in Aldous Huxley and Robert Ettinger.[196]

Edward Fredkin argues that "artificial intelligence is the next stage 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 book of the same name in 1998.[197]

Risks

Technological unemployment

Main articles: Workplace impact of artificial intelligence and Technological unemployment

In the past, technology has tended to increase rather than reduce total employment, but economists acknowledge that "we're in uncharted territory" with AI.[198] 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.[199] Subjective estimates of the risk vary widely; for example, Michael Osborne and Carl Benedikt Frey estimate 47% of U.S. jobs are at "high risk" of potential automation, while an OECD report classifies only 9% of U.S. jobs as "high risk".[t][201]

Unlike previous waves of automation, many middle-class jobs may be eliminated by artificial intelligence; *The Economist* states 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".[202] 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.[203]

Bad actors and weaponized AI

Main articles: Lethal autonomous weapon, Artificial intelligence arms race, and AI safety

AI provides a number of tools that are particularly useful for authoritarian governments: smart spyware, face recognition and voice recognition allow widespread surveillance; such surveillance allows machine learning to classify potential enemies of the state and can prevent them from hiding; recommendation systems can precisely target propaganda and misinformation for maximum effect; deepfakes aid in producing misinformation; advanced AI can make centralized decision making more competitive with liberal and decentralized systems such as markets.[204]

Terrorists, criminals and rogue states may use other forms of weaponized AI such as advanced digital warfare and lethal autonomous weapons. By 2015, over fifty countries were reported to be researching battlefield robots.[205]

Machine-learning AI is also able to design tens of thousands of toxic molecules in a matter of hours.[206]

Algorithmic bias

Main article: Algorithmic bias

AI programs can become biased after learning from real-world data. It is not typically introduced by the system designers but is learned by the program, and thus the programmers are often unaware that the bias exists.[207] Bias can be inadvertently introduced by the way training data is selected.[208] It can also emerge from correlations: AI is used to classify individuals into groups and then make predictions assuming that the individual will resemble other members of the group. In some cases, this assumption may be unfair.[209] An example of this is COMPAS, a commercial program widely used by U.S. courts to assess the likelihood of a defendant becoming a recidivist. ProPublica claims that the COMPAS-assigned recidivism risk level of black defendants is far more likely to be overestimated than that of white defendants, despite the fact that the program was not told the races of the defendants.[210]

Health equity issues may also be exacerbated when many-to-many mapping are done without taking steps to ensure equity for populations at risk for bias. At this time equity-focused tools and regulations are not in place to ensure equity application representation and usage.[211] Other examples where algorithmic bias can lead to unfair outcomes are when AI is used for credit rating or hiring.

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 recommending 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.[212]

Existential risk

Main articles: Existential risk from artificial general intelligence, AI alignment, and AI safety

Superintelligent AI may be able to improve itself to the point that humans could not control it. This could, as physicist Stephen Hawking puts it, "spell the end of the human race".[213] Philosopher Nick Bostrom argues that sufficiently intelligent AI, if it chooses actions based on achieving some goal, will exhibit convergent behavior such as acquiring resources or protecting itself from being shut down. If this AI's goals do not fully reflect humanity's, it might need to harm humanity to acquire more resources or prevent itself from being shut down, ultimately to better achieve its goal. He concludes that AI poses a risk to mankind, however humble or "friendly" its stated goals might be.[214] Political scientist Charles T. Rubin argues that "any sufficiently advanced benevolence may be indistinguishable from malevolence." Humans should not assume machines or robots would treat us favorably because there is no *a priori* reason to believe that they would share our system of morality.[215]

The opinion of experts and industry insiders is mixed, with sizable fractions both concerned and unconcerned by risk from eventual superhumanly-capable AI.[216] Stephen Hawking, Microsoft founder Bill Gates, history professor Yuval Noah Harari, and SpaceX founder Elon Musk have all expressed serious misgivings about the future of AI.[217] Prominent tech titans including Peter Thiel (Amazon Web Services) and Musk have committed more than $1 billion to nonprofit companies that champion responsible AI development, such as OpenAI and the Future of Life Institute.[218] Mark Zuckerberg (CEO, Facebook) has said that artificial intelligence is helpful in its current form and will continue to assist humans.[219] Other experts argue is that the risks are far enough in the future to not be worth researching, or that humans will be valuable from the perspective of a superintelligent machine.[220] Rodney Brooks, in particular, has said that "malevolent" AI is still centuries away.[u]

Copyright

AI's decisions making abilities raises the questions of legal responsibility and copyright status of created works. This issues are being refined in various jurisdictions.[222]

Ethical machines

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.[223]

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.[224] Machine ethics is also called machine morality, computational ethics or computational morality,[224] and was founded at an AAAI symposium in 2005.[225]

Other approaches include Wendell Wallach's "artificial moral agents"[226] and Stuart J. Russell's three principles for developing provably beneficial machines.[227]

Regulation

Main articles: Regulation of artificial intelligence, Regulation of algorithms, and AI safety

The regulation of artificial intelligence is the development of public sector policies and laws for promoting and regulating artificial intelligence (AI); it is therefore related to the broader regulation of algorithms.[228] The regulatory and policy landscape for AI is an emerging issue in jurisdictions globally.[229] Between 2016 and 2020, more than 30 countries adopted dedicated strategies for AI.[45] Most EU member states had released national AI strategies, as had Canada, China, India, Japan, Mauritius, the Russian Federation, Saudi Arabia, United Arab Emirates, US and Vietnam. Others were in the process of elaborating their own AI strategy, including Bangladesh, Malaysia and Tunisia.[45] 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.[45] Henry Kissinger, Eric Schmidt, and Daniel Huttenlocher published a joint statement in November 2021 calling for a government commission to regulate AI.[230]

In fiction

Main article: Artificial intelligence in fiction

Thought-capable artificial beings have appeared as storytelling devices since antiquity,[16] and have been a persistent theme in science fiction.[18]

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

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.[231]

Isaac Asimov introduced the Three Laws of Robotics in many books and stories, most notably the "Multivac" series about a super-intelligent computer of the same name. Asimov's laws are often brought up during lay discussions of machine ethics;[232] 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.[233]

Transhumanism (the merging of humans and machines) is explored in the manga *Ghost in the Shell* and the science-fiction series *Dune*.

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.[234]

See also

- AI safety – Research area on making AI safe and beneficial
- AI alignment – Conformance to the intended objective
- Artificial intelligence in healthcare - Machine-learning algorithms and software in the analysis, presentation, and comprehension of complex medical and health care data
- Artificial intelligence arms race – Arms race for the most advanced AI-related technologies
- Behavior selection algorithm – Algorithm that selects actions for intelligent agents
- Business process automation – Technology-enabled automation of complex business processes
- Case-based reasoning – Process of solving new problems based on the solutions of similar past problems
- Emergent algorithm – Algorithm exhibiting emergent behavior
- Female gendering of AI technologies – Design of digital assistants as female
- Glossary of artificial intelligence – List of definitions of terms and concepts commonly used in the study of artificial intelligence
- Operations research – Discipline concerning the application of advanced analytical methods
- Robotic process automation – Form of business process automation technology
- Synthetic intelligence – Alternate term for or form of artificial intelligence
- Universal basic income – Welfare system of unconditional income
- Weak artificial intelligence – Form of artificial intelligence
- Data sources – The list of data sources for study and research

Explanatory notes

a. ^ *a* *b* This list of intelligent traits is based on the topics covered by the major AI textbooks, including: Russell & Norvig (2003), Luger & Stubblefield (2004), Poole, Mackworth & Goebel (1998) and Nilsson (1998)

b. ^ This statement comes from the proposal for the Dartmouth workshop of 1956, which reads: "Every aspect of learning or any other feature of intelligence can be so precisely described that a machine can be made to simulate it."[11]

c. ^ Russel and Norvig note in the textbook Artificial Intelligence: A Modern Approach (4th ed.), section 1.5: "In the longer term, we face the difficult problem of controlling superintelligent AI systems that may evolve in unpredictable ways." while referring to computer scientists, philosophers, and technologists.

d. ^ Daniel Crevier wrote "the conference is generally recognized as the official birthdate of the new science."[24] Russell and Norvifg call the conference "the birth of artificial intelligence."[25]

e. ^ Russell and Norvig wrote "for the next 20 years the field would be dominated by these people and their students."[25]

f. ^ Russell and Norvig wrote "it was astonishing whenever a computer did anything kind of smartish".[27]

g. ^ 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.

h. ^ Embodied approaches to AI[37] were championed by Hans Moravec[38] and Rodney Brooks[39] and went by many names: Nouvelle AI,[39] Developmental robotics,[40] situated AI, behavior-based AI as well as others. A similar movement in cognitive science was the embodied mind thesis.

i. ^ Clark wrote: "After a half-decade of quiet breakthroughs in artificial intelligence, 2015 has been a landmark year. Computers are smarter and learning faster than ever."[8]

j. ^ Alan Turing discussed the centrality of learning as early as 1950, in his classic paper "Computing Machinery and Intelligence".[65] In 1956, at the original Dartmouth AI summer conference, Ray Solomonoff wrote a report on unsupervised probabilistic machine learning: "An Inductive Inference Machine".[66]

k. ^ This is a form of Tom Mitchell's widely quoted definition of machine learning: "A computer program is set to learn from an experience *E* with respect to some task *T* and some performance measure *P* if its performance on *T* as measured by *P* improves with experience *E*."[67]

l. ^ Alan Turing suggested in "Computing Machinery and Intelligence" that a "thinking machine" would need to be educated like a child.[65] Developmental robotics is a modern version of the idea.[40]

m. ^ 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.[106]

n. ^ Expectation-maximization, one of the most popular algorithms in machine learning, allows clustering in the presence of unknown latent variables.[108]

o. ^ The Smithsonian reports: "Pluribus has bested poker pros in a series of six-player no-limit Texas Hold'em games, reaching a milestone in artificial intelligence research. It is the first bot to beat humans in a complex multiplayer competition."[151]

p. ^ See Problem of other minds

q. ^ Nils Nilsson wrote in 1983: "Simply put, there is wide disagreement in the field about what AI is all about."[170]

r. ^ 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."[175]

s. ^ Searle presented this definition of "Strong AI" in 1999.[186] 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."[187] Strong AI is defined similarly by Russell and Norvig: "The assertion that machines could possibly act intelligently (or, perhaps better, act *as if* they were intelligent) is called the 'weak AI' hypothesis by philosophers, and the assertion that machines that do so are actually thinking (as opposed to simulating thinking) is called the 'strong AI' hypothesis."[182]

t. ^ See table 4; 9% is both the OECD average and the US average.[200]

References

1. ^ Google (2016).
2. ^ McCorduck (2004), p. 204.

3. ^ Schank (1991), p. 38.
4. ^ Crevier (1993), p. 109.
5. ^ a b c Funding initiatives in the early 80s: 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 (2003, p. 24)
   - NRC (1999, pp. 210–211)
   - Newquist (1994, pp. 235–248)
6. ^ a b First AI Winter, Lighthill report, Mansfield Amendment
   - Crevier (1993, pp. 115–117)
   - Russell & Norvig (2003, p. 22)
   - NRC (1999, pp. 212–213)
   - Howe (1994)
   - Newquist (1994, pp. 189–201)
7. ^ a b Second AI Winter:
   - McCorduck (2004, pp. 430–435)
   - Crevier (1993, pp. 209–210)
   - NRC (1999, pp. 214–216)
   - Newquist (1994, pp. 301–318)
8. ^ a b c d Clark (2015b).
9. ^ a b AI widely used in late 1990s:
   - Russell & Norvig (2003, p. 28)
   - Kurzweil (2005, p. 265)
   - NRC (1999, pp. 216–222)
   - Newquist (1994, pp. 189–201)
10. ^ a b Pennachin & Goertzel (2007); Roberts (2016)
11. ^ McCarthy et al. (1955).
12. ^ Newquist (1994), pp. 45–53.
13. ^ "s AI Overhyped in 2022? Getting the Truth About the True Power". Analytics Insight. 21 March 2022. Archived from the original on 10 March 2023. Retrieved 11 March 2023.
14. ^ Giles, Martin (13 September 2018). "Artificial intelligence is often overhyped—and here's why that's dangerous". MIT Technology. Archived from the original on 11 March 2023. Retrieved 11 March 2023.
15. ^ Basen, Ira (21 February 2020). "Is AI overhyped? Researchers weigh in on technology's promise and problems". Canadian Broadcasting Corporation. Archived from the original on 11 March 2023. Retrieved 11 March 2023.
16. ^ a b AI in myth:
   - McCorduck (2004, pp. 4–5)
   - Russell & Norvig (2003, p. 939)
17. ^ McCorduck (2004), pp. 17–25.
18. ^ a b McCorduck (2004), pp. 340–400.
19. ^ Berlinski (2000).
20. ^ AI's immediate precursors:
   - McCorduck (2004, pp. 51–107)
   - Crevier (1993, pp. 27–32)
   - Russell & Norvig (2003, pp. 15, 940)
   - Moravec (1988, p. 3)
21. ^ Russell & Norvig (2009), p. 16.
22. ^ Manyika 2022, p. 9.
23. ^ Manyika 2022, p. 10.
24. ^ Crevier (1993), pp. 47–49.
25. ^ a b Russell & Norvig (2003), p. 17.
26. ^ Dartmouth workshop:
   - Russell & Norvig (2003, p. 17)
   - McCorduck (2004, pp. 111–136)
   - NRC (1999, pp. 200–201)

   The proposal:
   - McCarthy et al. (1955)
27. ^ Russell & Norvig (2003), p. 18.
28. ^ Successful Symbolic AI programs:
   - McCorduck (2004, pp. 243–252)
   - Crevier (1993, pp. 52–107)
   - Moravec (1988, p. 9)
   - Russell & Norvig (2003, pp. 18–21)
29. ^ AI heavily funded in 1960s:
   - McCorduck (2004, p. 131)
   - Crevier (1993, pp. 51, 64–65)
   - NRC (1999, pp. 204–205)
30. ^ Howe (1994).
31. ^ Newquist (1994), pp. 86–86.
32. ^ Simon (1965, p. 96) quoted in Crevier (1993, p. 109)
33. ^ Minsky (1967, p. 2) quoted in Crevier (1993, p. 109)
34. ^ Lighthill (1973).
35. ^ Expert systems:
   - Russell & Norvig (2003, pp. 22–24)
   - Luger & Stubblefield (2004, pp. 227–331)
   - Nilsson (1998, chpt. 17.4)
   - McCorduck (2004, pp. 327–335, 434–435)
   - Crevier (1993, pp. 145–62, 197–203)
   - Newquist (1994, pp. 155–183)
36. ^ Nilsson (1998), p. 7.
37. ^ McCorduck (2004), pp. 454–462.
38. ^ Moravec (1988).
39. ^ a b Brooks (1990).
40. ^ a b Developmental robotics:
   - Weng et al. (2001)
   - Lungarella et al. (2003)
   - Asada et al. (2009)
   - Oudeyer (2010)
41. ^ Revival of connectionism:
   - Crevier (1993, pp. 214–215)
   - Russell & Norvig (2003, p. 25)
42. ^ Formal and narrow methods adopted in the 1990s:
   - Russell & Norvig (2003, pp. 25–26)
   - McCorduck (2004, pp. 486–487)
43. ^ McKinsey (2018).
44. ^ MIT Sloan Management Review (2018); Lorica (2017)
45. ^ a b c d UNESCO (2021).
46. ^ Lanier, Jaron (20 April 2023). "Annals of Artificial Intelligence - There Is No A.I. - There are ways of controlling the new technology—but first we have to stop mythologizing it". *The New Yorker*. Archived from the original on 23 April 2023. Retrieved 24 April 2023.
47. ^ Bogdan, Dennis (2 February 2023). "Comment - In the Age of A.I., Major in Being Human - David Brooks". *The New York Times*. Archived from the original on 3 February 2023. Retrieved 24 April 2023.
48. ^ Problem solving, puzzle solving, game playing and deduction:
   - Russell & Norvig (2003, chpt. 3–9)

- Poole, Mackworth & Goebel (1998, chpt. 2,3,7,9)
- Luger & Stubblefield (2004, chpt. 3,4,6,8)
- Nilsson (1998, chpt. 7–12)

49. ^ Uncertain reasoning:
    - Russell & Norvig (2003, pp. 452–644)
    - Poole, Mackworth & Goebel (1998, pp. 345–395)
    - Luger & Stubblefield (2004, pp. 333–381)
    - Nilsson (1998, chpt. 19)
50. ^ a b c Intractability and efficiency and the combinatorial explosion:
    - Russell & Norvig (2003, pp. 9, 21–22)
51. ^ a b c Psychological evidence of the prevalence sub-symbolic reasoning and knowledge:
    - Kahneman (2011)
    - Wason & Shapiro (1966)
    - Kahneman, Slovic & Tversky (1982)
    - Dreyfus & Dreyfus (1986)
52. ^ Knowledge representation and knowledge engineering:
    - Russell & Norvig (2003, pp. 260–266, 320–363)
    - 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)
53. ^ Russell & Norvig (2003), pp. 320–328.
54. ^ a b c Representing categories and relations: Semantic networks, description logics, inheritance (including frames and scripts):
    - Russell & Norvig (2003, pp. 349–354),
    - Poole, Mackworth & Goebel (1998, pp. 174–177),
    - Luger & Stubblefield (2004, pp. 248–258),
    - Nilsson (1998, chpt. 18.3)
55. ^ a b Representing events and time:Situation calculus, event calculus, fluent calculus (including solving the frame problem):
    - Russell & Norvig (2003, pp. 328–341),
    - Poole, Mackworth & Goebel (1998, pp. 281–298),
    - Nilsson (1998, chpt. 18.2)
56. ^ a b Causal calculus:
    - Poole, Mackworth & Goebel (1998, pp. 335–337)
57. ^ a b Representing knowledge about knowledge: Belief calculus, modal logics:
    - Russell & Norvig (2003, pp. 341–344),
    - Poole, Mackworth & Goebel (1998, pp. 275–277)
58. ^ a b Default reasoning, Frame problem, default logic, non-monotonic logics, circumscription, closed world assumption, abduction:
    - Russell & Norvig (2003, pp. 354–360)
    - 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").
59. ^ a b Breadth of commonsense knowledge:
    - Russell & Norvig (2003, p. 21),
    - Crevier (1993, pp. 113–114),
    - Moravec (1988, p. 13),
    - Lenat & Guha (1989, Introduction)
60. ^ Smoliar & Zhang (1994).
61. ^ Neumann & Möller (2008).
62. ^ Kuperman, Reichley & Bailey (2006).
63. ^ McGarry (2005).
64. ^ Bertini, Del Bimbo & Torniai (2006).
65. ^ a b Turing (1950).
66. ^ Solomonoff (1956).
67. ^ Russell & Norvig (2003), pp. 649–788.
68. ^ Learning:
    - Russell & Norvig (2003, pp. 649–788)
    - Poole, Mackworth & Goebel (1998, pp. 397–438)
    - Luger & Stubblefield (2004, pp. 385–542)
    - Nilsson (1998, chpt. 3.3, 10.3, 17.5, 20)
69. ^ Reinforcement learning:
    - Russell & Norvig (2003, pp. 763–788)
    - Luger & Stubblefield (2004, pp. 442–449)
70. ^ The Economist (2016).
71. ^ Jordan & Mitchell (2015).
72. ^ Natural language processing (NLP):
    - Russell & Norvig (2003, pp. 790–831)
    - Poole, Mackworth & Goebel (1998, pp. 91–104)
    - Luger & Stubblefield (2004, pp. 591–632)
73. ^ Applications of NLP:
    - Russell & Norvig (2003, pp. 840–857)
    - Luger & Stubblefield (2004, pp. 623–630)
74. ^ Modern statistical approaches to NLP:
    - Cambria & White (2014)
75. ^ Vincent (2019).
76. ^ Machine perception:
    - Russell & Norvig (2003, pp. 537–581, 863–898)
    - Nilsson (1998, ~chpt. 6)
77. ^ Speech recognition:
    - Russell & Norvig (2003, pp. 568–578)
78. ^ Object recognition:
    - Russell & Norvig (2003, pp. 885–892)
79. ^ Computer vision:
    - Russell & Norvig (2003, pp. 863–898)
    - Nilsson (1998, chpt. 6)
80. ^ MIT AIL (2014).
81. ^ Affective computing:
    - Thro (1993)
    - Edelson (1991)
    - Tao & Tan (2005)
    - Scassellati (2002)
82. ^ Waddell (2018).
83. ^ Poria et al. (2017).
84. ^ *The Society of Mind*:
    - Minsky (1986)

    Moravec's "golden spike":
    - Moravec (1988, p. 20)

    Multi-agent systems, hybrid intelligent systems, agent architectures, cognitive architecture:
    - Russell & Norvig (2003, pp. 27, 932, 970–972)

- Russell & Norvig (2003, pp. 27, 932, 970–972)
- Nilsson (1998, chpt. 25)

85. ^ Domingos (2015), Chpt. 9.
86. ^ Artificial brain as an approach to AGI:
    - Russell & Norvig (2003, p. 957)
    - Crevier (1993, pp. 271 & 279)
    - Goertzel *et al.* (2010)

    A few of the people who make some form of the argument:
    - Moravec (1988, p. 20)
    - Kurzweil (2005, p. 262)
    - Hawkins & Blakeslee (2005)
87. ^ Search algorithms:
    - Russell & Norvig (2003, pp. 59–189)
    - Poole, Mackworth & Goebel (1998, pp. 113–163)
    - Luger & Stubblefield (2004, pp. 79–164, 193–219)
    - Nilsson (1998, chpt. 7–12)
88. ^ Forward chaining, backward chaining, Horn clauses, and logical deduction as search:
    - Russell & Norvig (2003, pp. 217–225, 280–294)
    - Poole, Mackworth & Goebel (1998, pp. ~46–52)
    - Luger & Stubblefield (2004, pp. 62–73)
    - Nilsson (1998, chpt. 4.2, 7.2)
89. ^ State space search and planning:
    - Russell & Norvig (2003, pp. 382–387)
    - Poole, Mackworth & Goebel (1998, pp. 298–305)
    - Nilsson (1998, chpt. 10.1–2)
90. ^ Moving and configuration space:
    - Russell & Norvig (2003, pp. 916–932)
91. ^ Uninformed searches (breadth first search, depth-first search and general state space search):
    - Russell & Norvig (2003, pp. 59–93)
    - Poole, Mackworth & Goebel (1998, pp. 113–132)
    - Luger & Stubblefield (2004, pp. 79–121)
    - Nilsson (1998, chpt. 8)
92. ^ Heuristic or informed searches (e.g., greedy best first and A*):
    - Russell & Norvig (2003, pp. 94–109)
    - Poole, Mackworth & Goebel (1998, pp. pp. 132–147)
    - Poole & Mackworth (2017, Section 3.6)
    - Luger & Stubblefield (2004, pp. 133–150)
93. ^ Tecuci (2012).
94. ^ Optimization searches:
    - Russell & Norvig (2003, pp. 110–116, 120–129)
    - Poole, Mackworth & Goebel (1998, pp. 56–163)
    - Luger & Stubblefield (2004, pp. 127–133)
95. ^ Genetic programming and genetic algorithms:
    - Luger & Stubblefield (2004, pp. 509–530)
    - Nilsson (1998, chpt. 4.2)
96. ^ Artificial life and society based learning:
    - Luger & Stubblefield (2004, pp. 530–541)
    - Merkle & Middendorf (2013)
97. ^ Logic:
    - Russell & Norvig (2003, pp. 194–310),
    - Luger & Stubblefield (2004, pp. 35–77),
    - Nilsson (1998, chpt. 13–16)
98. ^ Satplan:
    - Russell & Norvig (2003, pp. 402–407),
    - Poole, Mackworth & Goebel (1998, pp. 300–301),
    - Nilsson (1998, chpt. 21)
99. ^ Explanation based learning, relevance based learning, inductive logic programming, case based reasoning:
    - Russell & Norvig (2003, pp. 678–710),
    - Poole, Mackworth & Goebel (1998, pp. 414–416),
    - Luger & Stubblefield (2004, pp. ~422–442),
    - Nilsson (1998, chpt. 10.3, 17.5)
100. ^ Propositional logic:
    - Russell & Norvig (2003, pp. 204–233),
    - Luger & Stubblefield (2004, pp. 45–50)
    - Nilsson (1998, chpt. 13)
101. ^ First-order logic and features such as equality:
    - Russell & Norvig (2003, pp. 240–310),
    - Poole, Mackworth & Goebel (1998, pp. 268–275),
    - Luger & Stubblefield (2004, pp. 50–62),
    - Nilsson (1998, chpt. 15)
102. ^ Fuzzy logic:
    - Russell & Norvig (2003, pp. 526–527)
    - Scientific American (1999)
103. ^ Abe, Jair Minoro; Nakamatsu, Kazumi (2009). "Multi-agent Systems and Paraconsistent Knowledge". *Knowledge Processing and Decision Making in Agent-Based Systems*. Studies in Computational Intelligence. Vol. 170. Springer Berlin Heidelberg. pp. 101–121. doi:10.1007/978-3-540-88049-3_5. eISSN 1860-9503. ISBN 978-3-540-88048-6. ISSN 1860-949X. Archived from the original on 9 February 2023. Retrieved 2 August 2022.
104. ^ Stochastic methods for uncertain reasoning:
    - Russell & Norvig (2003, pp. 462–644),
    - Poole, Mackworth & Goebel (1998, pp. 345–395),
    - Luger & Stubblefield (2004, pp. 165–191, 333–381),
    - Nilsson (1998, chpt. 19)
105. ^ Bayesian networks:
    - Russell & Norvig (2003, pp. 492–523),
    - Poole, Mackworth & Goebel (1998, pp. 361–381),
    - Luger & Stubblefield (2004, pp. ~182–190, ≈363–379),
    - Nilsson (1998, chpt. 19.3–4)
106. ^ Domingos (2015), chapter 6.
107. ^ Bayesian inference algorithm:
    - Russell & Norvig (2003, pp. 504–519),
    - Poole, Mackworth & Goebel (1998, pp. 361–381),
    - Luger & Stubblefield (2004, pp. ~363–379),
    - Nilsson (1998, chpt. 19.4 & 7)
108. ^ Domingos (2015), p. 210.
109. ^ Bayesian learning and the expectation-maximization algorithm:
    - Russell & Norvig (2003, pp. 712–724),
    - Poole, Mackworth & Goebel (1998, pp. 424–433),
    - Nilsson (1998, chpt. 20)
    - Domingos (2015, p. 210)
110. ^ Bayesian decision theory and Bayesian decision networks:

- Russell & Norvig (2003, pp. 597–600)

111. ^ a b c Stochastic temporal models:
    - Russell & Norvig (2003, pp. 537–581)

    Dynamic Bayesian networks:
    - Russell & Norvig (2003, pp. 551–557)

    Hidden Markov model:
    - (Russell & Norvig 2003, pp. 549–551)

    Kalman filters:
    - Russell & Norvig (2003, pp. 551–557)
112. ^ decision theory and decision analysis:
    - Russell & Norvig (2003, pp. 584–597),
    - Poole, Mackworth & Goebel (1998, pp. 381–394)
113. ^ Information value theory:
    - Russell & Norvig (2003, pp. 600–604)
114. ^ Markov decision processes and dynamic decision networks:
    - Russell & Norvig (2003, pp. 613–631)
115. ^ Game theory and mechanism design:
    - Russell & Norvig (2003, pp. 631–643)
116. ^ Statistical learning methods and classifiers:
    - Russell & Norvig (2003, pp. 712–754),
    - Luger & Stubblefield (2004, pp. 453–541)
117. ^ Decision tree:
    - Domingos (2015, p. 88)
    - Russell & Norvig (2003, pp. 653–664),
    - Poole, Mackworth & Goebel (1998, pp. 403–408),
    - Luger & Stubblefield (2004, pp. 408–417)
118. ^ K-nearest neighbor algorithm:
    - Domingos (2015, p. 187)
    - Russell & Norvig (2003, pp. 733–736)
119. ^ kernel methods such as the support vector machine:
    - Domingos (2015, p. 88)
    - Russell & Norvig (2003, pp. 749–752)

    Gaussian mixture model:
    - Russell & Norvig (2003, pp. 725–727)
120. ^ Domingos (2015), p. 152.
121. ^ Naive Bayes classifier:
    - Domingos (2015, p. 152)
    - Russell & Norvig (2003, p. 718)
122. ^ a b Neural networks:
    - Russell & Norvig (2003, pp. 736–748),
    - Poole, Mackworth & Goebel (1998, pp. 408–414),
    - Luger & Stubblefield (2004, pp. 453–505),
    - Nilsson (1998, chpt. 3)
    - Domingos (2015, Chapter 4)
123. ^ Classifier performance:
    - van der Walt & Bernard (2006)
    - Russell & Norvig (2009, 18.12: Learning from Examples: Summary)
124. ^ Backpropagation:
    - Russell & Norvig (2003, pp. 744–748),
    - Luger & Stubblefield (2004, pp. 467–474),
    - Nilsson (1998, chpt. 3.3)

    Paul Werbos' introduction of backpropagation to AI:
    - Werbos (1974); Werbos (1982)

    Automatic differentiation, an essential precursor:
    - Linnainmaa (1970); Griewank (2012)
125. ^ Competitive learning, Hebbian coincidence learning, Hopfield networks and attractor networks:
    - Luger & Stubblefield (2004, pp. 474–505)
126. ^ Feedforward neural networks, perceptrons and radial basis networks:
    - Russell & Norvig (2003, pp. 739–748, 758)
    - Luger & Stubblefield (2004, pp. 458–467)
127. ^ Schulz & Behnke (2012).
128. ^ Deep learning:
    - Goodfellow, Bengio & Courville (2016)
    - Hinton *et al.* (2016)
    - Schmidhuber (2015)
129. ^ Deng & Yu (2014), pp. 199–200.
130. ^ Ciresan, Meier & Schmidhuber (2012).
131. ^ Habibi (2017).
132. ^ Fukushima (2007).
133. ^ Recurrent neural networks, Hopfield nets:
    - Russell & Norvig (2003, p. 758)
    - Luger & Stubblefield (2004, pp. 474–505)
    - Schmidhuber (2015)
134. ^ Schmidhuber (2015).
135. ^ Werbos (1988); Robinson & Fallside (1987); Williams & Zipser (1994)
136. ^ Goodfellow, Bengio & Courville (2016); Hochreiter (1991)
137. ^ Hochreiter & Schmidhuber (1997); Gers, Schraudolph & Schraudolph (2002)
138. ^ Russell & Norvig (2009), p. 1.
139. ^ European Commission (2020), p. 1.
140. ^ CNN (2006).
141. ^ Targeted advertising:
    - Russell & Norvig (2009, p. 1)
    - Economist (2016)
    - Lohr (2016)
142. ^ Lohr (2016).
143. ^ Smith (2016).
144. ^ Rowinski (2013).
145. ^ Frangoul (2019).
146. ^ Brown (2019).
147. ^ "Artificial intelligence, immune to fear or favour, is helping to make China's foreign policy | South China Morning Post". 25 March 2023. Archived from the original on 25 March 2023. Retrieved 26 March 2023.
148. ^ McCorduck (2004), pp. 480–483.
149. ^ Markoff (2011).
150. ^ Google (2016); BBC (2016)
151. ^ Solly (2019).
152. ^ Bowling et al. (2015).
153. ^ Sample (2017).
154. ^ Anadiotis (2020).
155. ^ Heath (2020)

156. ^ Aletras *et al.* (2016).
157. ^ "AI Content Detector - Detect ChatGPT Plagiarism". *contentdetector.ai*. Archived from the original on 15 March 2023. Retrieved 16 March 2023.
158. ^ "Going Nowhere Fast? Smart Traffic Lights Can Help Ease Gridlock". 18 May 2022. Archived from the original on 22 December 2022. Retrieved 22 December 2022.
159. ^ "Intellectual Property and Frontier Technologies". *WIPO*. Archived from the original on 2 April 2022. Retrieved 30 March 2022.
160. ^ *a* *b* "WIPO Technology Trends 2019 – Artificial Intelligence" (PDF). *WIPO*. 2019. Archived (PDF) from the original on 9 October 2022.
161. ^ *a* *b* Turing (1950), p. 1.
162. ^ Turing's original publication of the Turing test in "Computing machinery and intelligence":
    - Turing (1950)

    Historical influence and philosophical implications:
    - Haugeland (1985, pp. 6–9)
    - Crevier (1993, p. 24)
    - McCorduck (2004, pp. 70–71)
    - Russell & Norvig (2021, pp. 2 and 984)
163. ^ Turing (1950), Under "The Argument from Consciousness".
164. ^ Russell & Norvig (2021), chpt. 2.
165. ^ Russell & Norvig (2021), p. 3.
166. ^ Maker (2006).
167. ^ McCarthy 1999.
168. ^ Minsky (1986).
169. ^ "Artificial intelligence - Google Search". *www.google.com*. Archived from the original on 1 December 2022. Retrieved 5 November 2022.
170. ^ Nilsson (1983), p. 10.
171. ^ Haugeland (1985), pp. 112–117.
172. ^ Physical symbol system hypothesis:
    - Newell & Simon (1976, p. 116)

    Historical significance:
    - McCorduck (2004, p. 153)
    - Russell & Norvig (2003, p. 18)
173. ^ Moravec's paradox:
    - Moravec (1988, pp. 15–16)
    - Minsky (1986, p. 29)
    - Pinker (2007, pp. 190–91)
174. ^ Dreyfus' critique of AI:
    - Dreyfus (1972)
    - Dreyfus & Dreyfus (1986)

    Historical significance and philosophical implications:
    - Crevier (1993, pp. 120–132)
    - McCorduck (2004, pp. 211–239)
    - Russell & Norvig (2003, pp. 950–952)
    - Fearn (2007, Chpt. 3)
175. ^ Crevier (1993), p. 125.
176. ^ Langley (2011).
177. ^ Katz (2012).
178. ^ Neats vs. scruffies, the historic debate:
    - McCorduck (2004, pp. 421–424, 486–489)
    - Crevier (1993, p. 168)
    - Nilsson (1983, pp. 10–11)

    A classic example of the "scruffy" approach to intelligence:
    - Minsky (1986)

    A modern example of neat AI and its aspirations:
    - Domingos (2015)
179. ^ Russell & Norvig (2003), pp. 25–26.
180. ^ Pennachin & Goertzel (2007).
181. ^ *a* *b* Roberts (2016).
182. ^ *a* *b* Russell & Norvig (2003), p. 947.
183. ^ Chalmers (1995).
184. ^ Dennett (1991).
185. ^ Horst (2005).
186. ^ Searle (1999).
187. ^ Searle (1980), p. 1.
188. ^ Searle's Chinese room argument:
    - Searle (1980). Searle's original presentation of the thought experiment.
    - Searle (1999).

    Discussion:
    - Russell & Norvig (2003, pp. 958–960)
    - McCorduck (2004, pp. 443–445)
    - Crevier (1993, pp. 269–271)
189. ^ Robot rights:
    - Russell & Norvig (2003, p. 964)
    - BBC (2006)
    - Maschafilm (2010) (the film Plug & Pray)
190. ^ Evans (2015).
191. ^ McCorduck (2004), pp. 19–25.
192. ^ Henderson (2007).
193. ^ Omohundro (2008).
194. ^ Vinge (1993).
195. ^ Russell & Norvig (2003), p. 963.
196. ^ Transhumanism:
    - Moravec (1988)
    - Kurzweil (2005)
    - Russell & Norvig (2003, p. 963)
197. ^ AI as evolution:
    - Edward Fredkin is quoted in McCorduck (2004, p. 401)
    - Butler (1863)
    - Dyson (1998)
198. ^ Ford & Colvin (2015); McGaughey (2018)
199. ^ IGM Chicago (2017).
200. ^ Arntz, Gregory & Zierahn (2016), p. 33.
201. ^ Lohr (2017); Frey & Osborne (2017); Arntz, Gregory & Zierahn (2016, p. 33)
202. ^ Morgenstern (2015).
203. ^ Mahdawi (2017); Thompson (2014)
204. ^ Harari (2018).
205. ^ Weaponized AI:
    - Robitzski (2018)
    - Sainato (2015)
206. ^ Urbina, Fabio; Lentzos, Filippa; Invernizzi, Cédric; Ekins, Sean (7 March 2022). "Dual use of artificial-intelligence-powered drug discovery". *Nature Machine Intelligence*. **4** (3): 189–191. doi:10.1038/s42256-022-00465-9. PMC 9544280. PMID 36211133. S2CID 247302391.
207. ^ CNA (2019).
208. ^ Goffrey (2008), p. 17.

209. ^ Lipartito (2011, p. 36); Goodman & Flaxman (2017, p. 6)
210. ^ Larson & Angwin (2016).
211. ^ Berdahl, Carl Thomas; Baker, Lawrence; Mann, Sean; Osoba, Osonde; Girosi, Federico (7 February 2023). "Strategies to Improve the Impact of Artificial Intelligence on Health Equity: Scoping Review". *JMIR AI*. **2**: e42936. doi:10.2196/42936. ISSN 2817-1705. S2CID 256681439. Archived from the original on 21 February 2023. Retrieved 21 February 2023.
212. ^ Dockrill, Peter, *Robots With Flawed AI Make Sexist And Racist Decisions, Experiment Shows Archived 27 June 2022 at the Wayback Machine*, Science Alert, 27 June 2022
213. ^ Cellan-Jones (2014).
214. ^ Bostrom (2014); Müller & Bostrom (2014); Bostrom (2015)

AI textbooks

These were the four the most widely used AI textbooks in 2008:

- Luger, George; Stubblefield, William (2004). *Artificial Intelligence: Structures and Strategies for Complex Problem Solving* (5th ed.). Benjamin/Cummings. ISBN 978-0-8053-4780-7. Archived from the original on 26 July 2020. Retrieved 17 December 2019.
- Nilsson, Nils (1998). *Artificial Intelligence: A New Synthesis*. Morgan Kaufmann. ISBN 978-1-55860-467-4. Archived from the original on 26 July 2020. Retrieved 18 November 2019.
- Russell, Stuart J.; Norvig, Peter (2003), *Artificial Intelligence: A Modern Approach* (2nd ed.), Upper Saddle River, New Jersey: Prentice Hall, ISBN 0-13-790395-2.
- Poole, David; Mackworth, Alan; Goebel, Randy (1998). *Computational Intelligence: A Logical Approach*. New York: Oxford University Press. ISBN 978-0-19-510270-3. Archived from the original on 26 July 2020. Retrieved 22 August 2020.

Later editions.

- Russell, Stuart J.; Norvig, Peter (2009). *Artificial Intelligence: A Modern Approach* (3rd ed.). Upper Saddle River, New Jersey: Prentice Hall. ISBN 978-0-13-604259-4..
- Poole, David; Mackworth, Alan (2017). *Artificial Intelligence: Foundations of Computational Agents* (2nd ed.). Cambridge University Press. ISBN 978-1-107-19539-4. Archived from the original on 7 December 2017. Retrieved 6 December 2017.

The two most widely used textbooks in 2021.Open Syllabus: Explorer Archived 7 October 2021 at the Wayback Machine

- Russell, Stuart J.; Norvig, Peter (2021). *Artificial Intelligence: A Modern Approach* (4th ed.). Hoboken: Pearson. ISBN 9780134610993. LCCN 20190474.
- Knight, Kevin; Rich, Elaine (1 January 2010). *Artificial Intelligence* (3rd ed.). Mc Graw Hill India. ISBN 9780070087705.

History of AI

- Crevier, Daniel (1993). *AI: The Tumultuous Search for Artificial Intelligence*. New York, NY: BasicBooks. ISBN 0-465-02997-3..
- McCorduck, Pamela (2004), *Machines Who Think* (2nd ed.), Natick, MA: A. K. Peters, Ltd., ISBN 1-56881-205-1.
- Newquist, HP (1994). *The Brain Makers: Genius, Ego, And Greed In The Quest For Machines That Think*. New York: Macmillan/SAMS. ISBN 978-0-672-30412-5.
- Nilsson, Nils (2009). *The Quest for Artificial Intelligence: A History of Ideas and Achievements*. New York: Cambridge University Press. ISBN 978-0-521-12293-1.

Other sources

- McCarthy, John (1999), *What is AI?*, archived from the original on 4 December 2022, retrieved 4 December 2022
- Werbos, P. J. (1988), "Generalization of backpropagation with application to a recurrent gas market model", *Neural Networks*, **1** (4): 339–356, doi:10.1016/0893-6080(88)90007-X, archived from the original on 29 October 2021, retrieved 29 September 2021
- Gers, Felix A.; Schraudolph, Nicol N.; Schraudolph, Jürgen (2002). "Learning Precise Timing with LSTM Recurrent Networks" (PDF). *Journal of Machine Learning Research*. **3**: 115–143. Archived (PDF) from the original on 9 October 2022. Retrieved 13 June 2017.
- Deng, L.; Yu, D. (2014). "Deep Learning: Methods and Applications" (PDF). *Foundations and Trends in Signal Processing*. **7** (3–4): 1–199. doi:10.1561/2000000039. Archived (PDF) from the original on 14 March 2016. Retrieved 18 October 2014.
- Schulz, Hannes; Behnke, Sven (1 November 2012). "Deep Learning". *KI – Künstliche Intelligenz*. **26** (4): 357–363. doi:10.1007/s13218-012-0198-z. ISSN 1610-1987. S2CID 220523562.
- Fukushima, K. (2007). "Neocognitron". *Scholarpedia*. **2** (1): 1717. Bibcode:2007SchpJ...2.1717F. doi:10.4249/scholarpedia.1717. was introduced by Kunihiko Fukushima in 1980.
- Habibi, Aghdam, Hamed (30 May 2017). *Guide to convolutional neural networks : a practical application to traffic-sign detection and classification*. Heravi, Elnaz Jahani. Cham, Switzerland. ISBN 9783319575490. OCLC 987790957.
- Ciresan, D.; Meier, U.; Schmidhuber, J. (2012). "Multi-column deep neural networks for image classification". *2012 IEEE Conference on Computer Vision and Pattern Recognition*. pp. 3642–3649. arXiv:1202.2745. doi:10.1109/cvpr.2012.6248110. ISBN 978-1-4673-1228-8. S2CID 2161592.
- "From not working to neural networking". *The Economist*. 2016. Archived from the original on 31 December 2016. Retrieved 26 April 2018.
- Thompson, Derek (23 January 2014). "What Jobs Will the Robots Take?". *The Atlantic*. Archived from the original on 24 April 2018. Retrieved 24 April 2018.
- Scassellati, Brian (2002). "Theory of mind for a humanoid robot". *Autonomous Robots*. **12** (1): 13–24. doi:10.1023/A:1013298507114. S2CID 1979315.
- Sample, Ian (14 March 2017). "Google's DeepMind makes AI program that can learn like a human". *The Guardian*. Archived from the original on 26 April 2018. Retrieved 26 April 2018.
- Heath, Nick (11 December 2020). "What is AI? Everything you need to know about Artificial Intelligence". *ZDNet*. Archived from the original on 2 March 2021. Retrieved 1 March 2021.
- Bowling, Michael; Burch, Neil; Johanson, Michael; Tammelin, Oskari (9 January 2015). "Heads-up limit hold'em poker is solved". *Science*. **347** (6218): 145–149. Bibcode:2015Sci...347..145B. doi:10.1126/science.1259433. ISSN 0036-8075. PMID 25574016. S2CID 3796371. Archived from the original on 1 August 2022. Retrieved 30 June 2022.
- Solly, Meilan (15 July 2019). "This Poker-Playing A.I. Knows When to Hold 'Em and When to Fold 'Em". *Smithsonian*. Archived from the original on 26 September 2021. Retrieved 1 October 2021.
- "Artificial intelligence: Google's AlphaGo beats Go master Lee Se-dol". *BBC News*. 12 March 2016. Archived from the original on 26 August 2016. Retrieved 1 October 2016.
- Rowinski, Dan (15 January 2013). "Virtual Personal Assistants & The Future Of Your Smartphone [Infographic]". *ReadWrite*. Archived from the original on 22 December 2015.
- Manyika, James (2022). "Getting AI Right: Introductory Notes on AI & Society". *Daedalus*. **151** (2): 5–27. doi:10.1162/daed_e_01897. S2CID 248377878. Archived from the original on 5 May 2022. Retrieved 5 May 2022.
- Markoff, John (16 February 2011). "Computer Wins on 'Jeopardy!': Trivial, It's Not". *The New York Times*. Archived from the original on 22 October 2014. Retrieved 25 October 2014.
- Anadiotis, George (1 October 2020). "The state of AI in 2020: Democratization, industrialization, and the way to artificial general intelligence". *ZDNet*. Archived from the original on 15 March 2021. Retrieved 1 March 2021.
- Goertzel, Ben; Lian, Ruiting; Arel, Itamar; de Garis, Hugo; Chen, Shuo (December 2010). "A world survey of artificial brain projects, Part II: Biologically inspired cognitive architectures". *Neurocomputing*. **74** (1–3): 30–49. doi:10.1016/j.neucom.2010.08.012.
- Robinson, A. J.; Fallside, F. (1987), "The utility driven dynamic error propagation network.", *Technical Report CUED/F-INFENG/TR.1*, Cambridge University Engineering Department
- Hochreiter, Sepp (1991). *Untersuchungen zu dynamischen neuronalen Netzen* (PDF) (diploma thesis). Munich: Institut f. Informatik, Technische Univ. Archived from the original (PDF) on 6 March 2015. Retrieved 16 April 2016.
- Williams, R. J.; Zipser, D. (1994), "Gradient-based learning algorithms for recurrent networks and their computational complexity", *Back-propagation: Theory, Architectures and Applications*, Hillsdale, NJ: Erlbaum
- Hochreiter, Sepp; Schmidhuber, Jürgen (1997), "Long Short-Term Memory", *Neural Computation*, **9** (8): 1735–1780, doi:10.1162/neco.1997.9.8.1735, PMID 9377276, S2CID 1915014
- Goodfellow, Ian; Bengio, Yoshua; Courville, Aaron (2016), *Deep Learning*, MIT Press., archived from the original on 16 April 2016, retrieved 12 November 2017
- Hinton, G.; Deng, L.; Yu, D.; Dahl, G.; Mohamed, A.; Jaitly, N.; Senior, A.; Vanhoucke, V.; Nguyen, P.; Sainath, T.; Kingsbury, B. (2012). "Deep Neural Networks for Acoustic Modeling in Speech Recognition – The shared views of four research groups". *IEEE Signal Processing Magazine*. **29** (6): 82–97. Bibcode:2012ISPM...29...82H. doi:10.1109/msp.2012.2205597. S2CID 206485943.
- Schmidhuber, J. (2015). "Deep Learning in Neural Networks: An Overview". *Neural Networks*. **61**: 85–117. arXiv:1404.7828. doi:10.1016/j.neunet.2014.09.003. PMID 25462637. S2CID 11715509.
- Linnainmaa, Seppo (1970). *The representation of the cumulative rounding error of an algorithm as a Taylor expansion of the local rounding errors* (Thesis) (in Finnish). Univ. Helsinki, 6–7.|
- Griewank, Andreas (2012). "Who Invented the Reverse Mode of Differentiation? Optimization Stories". *Documenta Matematica, Extra Volume ISMP*: 389–400.
- Werbos, Paul (1974). *Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences* (Ph.D. thesis). Harvard University.
- Werbos, Paul (1982). "Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences" (PDF). *System Modeling and Optimization*. Applications of advances in nonlinear sensitivity analysis. Berlin, Heidelberg: Springer. Archived from the original (PDF) on 14 April 2016. Retrieved 16 April 2016.
- "What is 'fuzzy logic'? Are there computers that are inherently fuzzy and do not apply the usual binary logic?". *Scientific American*. 21 October 1999. Archived from the original on 6 May 2018. Retrieved 5 May 2018.
- Merkle, Daniel; Middendorf, Martin (2013). "Swarm Intelligence". In Burke, Edmund K.; Kendall, Graham (eds.). *Search Methodologies: Introductory Tutorials in Optimization and Decision Support Techniques*. Springer Science & Business Media. ISBN 978-1-4614-6940-7.
- van der Walt, Christiaan; Bernard, Etienne (2006). "Data characteristics that determine classifier performance" (PDF). Archived from the original (PDF) on 25 March 2009. Retrieved 5 August 2009.
- Hutter, Marcus (2005). *Universal Artificial Intelligence*. Berlin: Springer. ISBN 978-3-540-22139-5.
- Howe, J. (November 1994). "Artificial Intelligence at Edinburgh University: a Perspective". Archived from the original on 15 May 2007. Retrieved 30 August 2007.
- Galvan, Jill (1 January 1997). "Entering the Posthuman Collective in Philip K. Dick's "Do Androids Dream of Electric Sheep?"". *Science Fiction Studies*. **24** (3): 413–429. JSTOR 4240644.
- McCauley, Lee (2007). "AI armageddon and the three laws of robotics". *Ethics and Information Technology*. **9** (2): 153–164. CiteSeerX 10.1.1.85.8904. doi:10.1007/s10676-007-9138-2. S2CID 37272949.
- Buttazzo, G. (July 2001). "Artificial consciousness: Utopia or real possibility?". *Computer*. **34** (7): 24–30. doi:10.1109/2.933500.
- Anderson, Susan Leigh (2008). "Asimov's "three laws of robotics" and machine metaethics". *AI & Society*. **22** (4): 477–493. doi:10.1007/s00146-007-0094-5. S2CID 1809459.
- Yudkowsky, E (2008), "Artificial Intelligence as a Positive and Negative Factor in Global Risk" (PDF), *Global Catastrophic Risks*, Oxford University Press, 2008, Bibcode:2008gcr..book..303Y, archived (PDF) from the original on 19 October 2013, retrieved 24 September 2021
- McGaughey, E (2018), *Will Robots Automate Your Job Away? Full Employment, Basic Income, and Economic Democracy*, p. SSRN part 2(3), SSRN 3044448, archived from the original on 24 May 2018, retrieved 12 January 2018
- IGM Chicago (30 June 2017). "Robots and Artificial Intelligence". *www.igmchicago.org*. Archived from the original on 1 May 2019. Retrieved 3 July 2019.
- Lohr, Steve (2017). "Robots Will Take Jobs, but Not as Fast as Some Fear, New Report Says". *The New York Times*. Archived from the original on 14 January 2018. Retrieved 13 January 2018.
- Frey, Carl Benedikt; Osborne, Michael A (1 January 2017). "The future of employment: How susceptible are jobs to computerisation?". *Technological Forecasting and Social Change*. **114**: 254–280. CiteSeerX 10.1.1.395.416. doi:10.1016/j.techfore.2016.08.019. ISSN 0040-1625.
- Arntz, Melanie; Gregory, Terry; Zierahn, Ulrich (2016), "The risk of automation for jobs in OECD countries: A comparative analysis", *OECD Social, Employment, and Migration Working Papers 189*
- Morgenstern, Michael (9 May 2015). "Automation and anxiety". *The Economist*. Archived from the original on 12 January 2018. Retrieved 13 January 2018.
- Mahdawi, Arwa (26 June 2017). "What jobs will still be around in 20 years? Read this to prepare your future". *The Guardian*. Archived from the original on 14 January 2018. Retrieved 13 January 2018.
- Rubin, Charles (Spring 2003). "Artificial Intelligence and Human Nature". *The New Atlantis*. **1**: 88–100. Archived from the original on 11 June 2012.
- Bostrom, Nick (2014). *Superintelligence: Paths, Dangers, Strategies*. Oxford University Press.
- Brooks, Rodney (10 November 2014). "artificial intelligence is a tool, not a threat". Archived from the original on 12 November 2014.
- Sainato, Michael (19 August 2015). "Stephen Hawking, Elon Musk, and Bill Gates Warn About Artificial Intelligence". *Observer*. Archived from the original on 30 October 2015. Retrieved 30 October 2015.
- Harari, Yuval Noah (October 2018). "Why Technology Favors Tyranny". *The Atlantic*. Archived from the original on 25 September 2021. Retrieved 23 September 2021.
- Robitzski, Dan (5 September 2018). "Five experts share what scares them the most about AI". Archived from the original on 8 December 2019. Retrieved 8 December 2019.

- Goffrey, Andrew (2008). "Algorithm". In Fuller, Matthew (ed.). *Software studies: a lexicon*. Cambridge, Mass.: MIT Press. pp. 15–20. ISBN 978-1-4356-4787-9.
- Lipartito, Kenneth (6 January 2011), *The Narrative and the Algorithm: Genres of Credit Reporting from the Nineteenth Century to Today* (PDF) (Unpublished manuscript), doi:10.2139/ssrn.1736283, S2CID 166742927, archived (PDF) from the original on 9 October 2022
- Goodman, Bryce; Flaxman, Seth (2017). "EU regulations on algorithmic decision-making and a "right to explanation"". *AI Magazine*. **38** (3): 50. arXiv:1606.08813. doi:10.1609/aimag.v38i3.2741. S2CID 7373959.
- CNA (12 January 2019). "Commentary: Bad news. Artificial intelligence is biased". *CNA*. Archived from the original on 12 January 2019. Retrieved 19 June 2020.
- Larson, Jeff; Angwin, Julia (23 May 2016). "How We Analyzed the COMPAS Recidivism Algorithm". *ProPublica*. Archived from the original on 29 April 2019. Retrieved 19 June 2020.
- Müller, Vincent C.; Bostrom, Nick (2014). "Future Progress in Artificial Intelligence: A Poll Among Experts" (PDF). *AI Matters*. **1** (1): 9–11. doi:10.1145/2639475.2639478. S2CID 8510016. Archived (PDF) from the original on 15 January 2016.
- Cellan-Jones, Rory (2 December 2014). "Stephen Hawking warns artificial intelligence could end mankind". *BBC News*. Archived from the original on 30 October 2015. Retrieved 30 October 2015.
- Rawlinson, Kevin (29 January 2015). "Microsoft's Bill Gates insists AI is a threat". *BBC News*. Archived from the original on 29 January 2015. Retrieved 30 January 2015.
- Holley, Peter (28 January 2015). "Bill Gates on dangers of artificial intelligence: 'I don't understand why some people are not concerned'". *The Washington Post*. ISSN 0190-8286. Archived from the original on 30 October 2015. Retrieved 30 October 2015.
- Gibbs, Samuel (27 October 2014). "Elon Musk: artificial intelligence is our biggest existential threat". *The Guardian*. Archived from the original on 30 October 2015. Retrieved 30 October 2015.
- Churm, Philip Andrew (14 May 2019). "Yuval Noah Harari talks politics, technology and migration". *euronews*. Archived from the original on 14 May 2019. Retrieved 15 November 2020.
- Bostrom, Nick (2015). "What happens when our computers get smarter than we are?". TED (conference). Archived from the original on 25 July 2020. Retrieved 30 January 2020.
- Post, Washington (2015). "Tech titans like Elon Musk are spending $1 billion to save you from terminators". *Chicago Tribune*. Archived from the original on 7 June 2016.
- Del Prado, Guia Marie (9 October 2015). "The mysterious artificial intelligence company Elon Musk invested in is developing game-changing smart computers". *Tech Insider*. Archived from the original on 30 October 2015. Retrieved 30 October 2015.
- FastCompany (15 January 2015). "Elon Musk Is Donating $10M Of His Own Money To Artificial Intelligence Research". *Fast Company*. Archived from the original on 30 October 2015. Retrieved 30 October 2015.
- Thibodeau, Patrick (25 March 2019). "Oracle CEO Mark Hurd sees no reason to fear ERP AI". *SearchERP*. Archived from the original on 6 May 2019. Retrieved 6 May 2019.
- Bhardwaj, Prachi (24 May 2018). "Mark Zuckerberg responds to Elon Musk's paranoia about AI: 'AI is going to... help keep our communities safe.'". *Business Insider*. Archived from the original on 6 May 2019. Retrieved 6 May 2019.
- Geist, Edward Moore (9 August 2015). "Is artificial intelligence really an existential threat to humanity?". *Bulletin of the Atomic Scientists*. Archived from the original on 30 October 2015. Retrieved 30 October 2015.
- Madrigal, Alexis C. (27 February 2015). "The case against killer robots, from a guy actually working on artificial intelligence". *Fusion.net*. Archived from the original on 4 February 2016. Retrieved 31 January 2016.
- Lee, Timothy B. (22 August 2014). "Will artificial intelligence destroy humanity? Here are 5 reasons not to worry". *Vox*. Archived from the original on 30 October 2015. Retrieved 30 October 2015.
- Law Library of Congress (U.S.). Global Legal Research Directorate, issuing body. (2019). *Regulation of artificial intelligence in selected jurisdictions*. LCCN 2019668143. OCLC 1110727808.
- Berryhill, Jamie; Heang, Kévin Kok; Clogher, Rob; McBride, Keegan (2019). *Hello, World: Artificial Intelligence and its Use in the Public Sector* (PDF). Paris: OECD Observatory of Public Sector Innovation. Archived (PDF) from the original on 20 December 2019. Retrieved 9 August 2020.
- Barfield, Woodrow; Pagallo, Ugo (2018). *Research handbook on the law of artificial intelligence*. Cheltenham, UK. ISBN 978-1-78643-904-8. OCLC 1039480085.
- Iphofen, Ron; Kritikos, Mihalis (3 January 2019). "Regulating artificial intelligence and robotics: ethics by design in a digital society". *Contemporary Social Science*. **16** (2): 170–184. doi:10.1080/21582041.2018.1563803. ISSN 2158-2041. S2CID 59298502.
- Wirtz, Bernd W.; Weyerer, Jan C.; Geyer, Carolin (24 July 2018). "Artificial Intelligence and the Public Sector – Applications and Challenges". *International Journal of Public Administration*. **42** (7): 596–615. doi:10.1080/01900692.2018.1498103. ISSN 0190-0692. S2CID 158829602. Archived from the original on 18 August 2020. Retrieved 22 August 2020.
- Buiten, Miriam C (2019). "Towards Intelligent Regulation of Artificial Intelligence". *European Journal of Risk Regulation*. **10** (1): 41–59. doi:10.1017/err.2019.8. ISSN 1867-299X.
- Wallach, Wendell (2010). *Moral Machines*. Oxford University Press.
- Brown, Eileen (5 November 2019). "Half of Americans do not believe deepfake news could target them online". *ZDNet*. Archived from the original on 6 November 2019. Retrieved 3 December 2019.
- Frangoul, Anmar (14 June 2019). "A Californian business is using A.I. to change the way we think about energy storage". *CNBC*. Archived from the original on 25 July 2020. Retrieved 5 November 2019.
- "The Economist Explains: Why firms are piling into artificial intelligence". *The Economist*. 31 March 2016. Archived from the original on 8 May 2016. Retrieved 19 May 2016.
- Lohr, Steve (28 February 2016). "The Promise of Artificial Intelligence Unfolds in Small Steps". *The New York Times*. Archived from the original on 29 February 2016. Retrieved 29 February 2016.
- Smith, Mark (22 July 2016). "So you think you chose to read this article?". *BBC News*. Archived from the original on 25 July 2016.
- Aletras, N.; Tsarapatsanis, D.; Preotiuc-Pietro, D.; Lampos, V. (2016). "Predicting judicial decisions of the European Court of Human Rights: a Natural Language Processing perspective". *PeerJ Computer Science*. **2**: e93. doi:10.7717/peerj-cs.93.
- Cadena, Cesar; Carlone, Luca; Carrillo, Henry; Latif, Yasir; Scaramuzza, Davide; Neira, Jose; Reid, Ian; Leonard, John J. (December 2016). "Past, Present, and Future of Simultaneous Localization and Mapping: Toward the Robust-Perception Age". *IEEE Transactions on Robotics*. **32** (6): 1309–1332. arXiv:1606.05830. doi:10.1109/TRO.2016.2624754. S2CID 2596787.
- Cambria, Erik; White, Bebo (May 2014). "Jumping NLP Curves: A Review of Natural Language Processing Research [Review Article]". *IEEE Computational Intelligence Magazine*. **9** (2): 48–57. doi:10.1109/MCI.2014.2307227. S2CID 206451986.
- Vincent, James (7 November 2019). "OpenAI has published the text-generating AI it said was too dangerous to share". *The Verge*. Archived from the original on 11 June 2020. Retrieved 11 June 2020.
- Jordan, M. I.; Mitchell, T. M. (16 July 2015). "Machine learning: Trends, perspectives, and prospects". *Science*. **349** (6245): 255–260. Bibcode:2015Sci...349..255J. doi:10.1126/science.aaa8415. PMID 26185243. S2CID 677218.
- Maschafilm (2010). "Content: Plug & Pray Film – Artificial Intelligence – Robots -". *plugandpray-film.de*. Archived from the original on 12 February 2016.
- Evans, Woody (2015). "Posthuman Rights: Dimensions of Transhuman Worlds". *Teknokultura*. **12** (2). doi:10.5209/rev_TK.2015.v12.n2.49072.
- Waddell, Kaveh (2018). "Chatbots Have Entered the Uncanny Valley". *The Atlantic*. Archived from the original on 24 April 2018. Retrieved 24 April 2018.
- Poria, Soujanya; Cambria, Erik; Bajpai, Rajiv; Hussain, Amir (September 2017). "A review of affective computing: From unimodal analysis to multimodal fusion". *Information Fusion*. **37**: 98–125. doi:10.1016/j.inffus.2017.02.003. hdl:1893/25490. S2CID 205433041. Archived from the original on 23 March 2023. Retrieved 27 April 2021.
- "Robots could demand legal rights". *BBC News*. 21 December 2006. Archived from the original on 15 October 2019. Retrieved 3 February 2011.
- Horst, Steven (2005). "The Computational Theory of Mind". *The Stanford Encyclopedia of Philosophy*. Archived from the original on 6 March 2016. Retrieved 7 March 2016.
- Omohundro, Steve (2008). *The Nature of Self-Improving Artificial Intelligence*. presented and distributed at the 2007 Singularity Summit, San Francisco, CA.
- Ford, Martin; Colvin, Geoff (6 September 2015). "Will robots create more jobs than they destroy?". *The Guardian*. Archived from the original on 16 June 2018. Retrieved 13 January 2018.
- *White Paper: On Artificial Intelligence – A European approach to excellence and trust* (PDF). Brussels: European Commission. 2020. Archived (PDF) from the original on 20 February 2020. Retrieved 20 February 2020.
- Anderson, Michael; Anderson, Susan Leigh (2011). *Machine Ethics*. Cambridge University Press.
- "Machine Ethics". *aaai.org*. Archived from the original on 29 November 2014.
- Russell, Stuart (8 October 2019). *Human Compatible: Artificial Intelligence and the Problem of Control*. United States: Viking. ISBN 978-0-525-55861-3. OCLC 1083694322.
- "AI set to exceed human brain power". *CNN*. 9 August 2006. Archived from the original on 19 February 2008.
- "Robots could demand legal rights". *BBC News*. 21 December 2006. Archived from the original on 15 October 2019. Retrieved 3 February 2011.
- "Kismet". MIT Artificial Intelligence Laboratory, Humanoid Robotics Group. Archived from the original on 17 October 2014. Retrieved 25 October 2014.
- Smoliar, Stephen W.; Zhang, HongJiang (1994). "Content based video indexing and retrieval". *IEEE MultiMedia*. **1** (2): 62–72. doi:10.1109/93.311653. S2CID 32710913.
- Neumann, Bernd; Möller, Ralf (January 2008). "On scene interpretation with description logics". *Image and Vision Computing*. **26** (1): 82–101. doi:10.1016/j.imavis.2007.08.013. S2CID 10767011.
- Kuperman, G. J.; Reichley, R. M.; Bailey, T. C. (1 July 2006). "Using Commercial Knowledge Bases for Clinical Decision Support: Opportunities, Hurdles, and Recommendations". *Journal of the American Medical Informatics Association*. **13** (4): 369–371. doi:10.1197/jamia.M2055. PMC 1513681. PMID 16622160.
- McGarry, Ken (1 December 2005). "A survey of interestingness measures for knowledge discovery". *The Knowledge Engineering Review*. **20** (1): 39–61. doi:10.1017/S0269888905000408. S2CID 14987656.
- Bertini, M; Del Bimbo, A; Torniai, C (2006). "Automatic annotation and semantic retrieval of video sequences using multimedia ontologies". *MM '06 Proceedings of the 14th ACM international conference on Multimedia*. 14th ACM international conference on Multimedia. Santa Barbara: ACM. pp. 679–682.
- Kahneman, Daniel (25 October 2011). *Thinking, Fast and Slow*. Macmillan. ISBN 978-1-4299-6935-2. Archived from the original on 15 March 2023. Retrieved 8 April 2012.
- Turing, Alan (1948), "Machine Intelligence", in Copeland, B. Jack (ed.), *The Essential Turing: The ideas that gave birth to the computer age*, Oxford: Oxford University Press, p. 412, ISBN 978-0-19-825080-7
- Domingos, Pedro (22 September 2015). *The Master Algorithm: How the Quest for the Ultimate Learning Machine Will Remake Our World*. Basic Books. ISBN 978-0465065707.
- Minsky, Marvin (1986), *The Society of Mind*, Simon and Schuster
- Pinker, Steven (4 September 2007) [1994], *The Language Instinct*, Perennial Modern Classics, Harper, ISBN 978-0-06-133646-1
- Chalmers, David (1995). "Facing up to the problem of consciousness". *Journal of Consciousness Studies*. **2** (3): 200–219. Archived from the original on 8 March 2005. Retrieved 11 October 2018.
- Roberts, Jacob (2016). "Thinking Machines: The Search for Artificial Intelligence". *Distillations*. Vol. 2, no. 2. pp. 14–23. Archived from the original on 19 August 2018. Retrieved 20 March 2018.
- Pennachin, C.; Goertzel, B. (2007). "Contemporary Approaches to Artificial General Intelligence". *Artificial General Intelligence*. Cognitive Technologies. Berlin, Heidelberg: Springer. doi:10.1007/978-3-540-68677-4_1. ISBN 978-3-540-23733-4.
- "Ask the AI experts: What's driving today's progress in AI?". *McKinsey & Company*. Archived from the original on 13 April 2018. Retrieved 13 April 2018.
- 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 19 May 2018. Retrieved 2 May 2018.
- Lorica, Ben (18 December 2017). "The state of AI adoption". *O'Reilly Media*. Archived from the original on 2 May 2018. Retrieved 2 May 2018.
- "AlphaGo – Google DeepMind". Archived from the original on 20 October 2021.
- Asada, M.; Hosoda, K.; Kuniyoshi, Y.; Ishiguro, H.; Inui, T.; Yoshikawa, Y.; Ogino, M.; Yoshida, C. (2009). "Cognitive developmental robotics: a survey". *IEEE Transactions on Autonomous Mental Development*. **1** (1): 12–34. doi:10.1109/tamd.2009.2021702. S2CID 10168773.
- Berlinski, David (2000). *The Advent of the Algorithm*. Harcourt Books. ISBN 978-0-15-601391-8. OCLC 46890682. Archived from the original on 26 July 2020. Retrieved 22 August 2020.
- Brooks, Rodney (1990). "Elephants Don't Play Chess" (PDF). *Robotics and Autonomous Systems*. **6** (1–2): 3–15. CiteSeerX 10.1.1.588.7539. doi:10.1016/S0921-8890(05)80025-9. Archived (PDF) from the original on 9 August 2007.
- Butler, Samuel (13 June 1863). "Darwin among the Machines". Letters to the Editor. *The Press*. Christchurch, New Zealand. Archived from the original on 19 September 2008. Retrieved 16 October 2014 – via Victoria University of Wellington.
- Clark, Jack (2015a). "Musk-Backed Group Probes Risks Behind Artificial Intelligence". *Bloomberg.com*. Archived from the original on 30 October 2015. Retrieved 30 October 2015.
- Clark, Jack (2015b). "Why 2015 Was a Breakthrough Year in Artificial Intelligence". *Bloomberg.com*. Archived from the original on 23 November 2016. Retrieved 23 November 2016.
- Dennett, Daniel (1991). *Consciousness Explained*. The Penguin Press. ISBN 978-0-7139-9037-9.
- Dreyfus, Hubert (1972). *What Computers Can't Do*. New York: MIT Press. ISBN 978-0-06-011082-6.
- Dreyfus, Hubert; Dreyfus, Stuart (1986). *Mind over Machine: The Power of Human Intuition and Expertise in the Era of the Computer*. Oxford, UK: Blackwell. ISBN 978-0-02-908060-3. Archived from the original on 26 July 2020. Retrieved 22 August 2020.
- Dyson, George (1998). *Darwin among the Machines*. Allan Lane Science. ISBN 978-0-7382-0030-9. Archived from the original on 26 July 2020. Retrieved 22 August 2020.
- Edelson, Edward (1991). *The Nervous System*. New York: Chelsea House. ISBN 978-0-7910-0464-7. Archived from the original on 26 July 2020. Retrieved 18 November 2019.
- Fearn, Nicholas (2007). *The Latest Answers to the Oldest Questions: A Philosophical Adventure with the World's Greatest Thinkers*. New York: Grove Press. ISBN 978-0-8021-1839-4.

- Haugeland, John (1985). *Artificial Intelligence: The Very Idea*. Cambridge, Mass.: MIT Press. ISBN 978-0-262-08153-5.
- Hawkins, Jeff; Blakeslee, Sandra (2005). *On Intelligence*. New York: Owl Books. ISBN 978-0-8050-7853-4.
- Henderson, Mark (24 April 2007). "Human rights for robots? We're getting carried away". *The Times Online*. London. Archived from the original on 31 May 2014. Retrieved 31 May 2014.
- Kahneman, Daniel; Slovic, D.; Tversky, Amos (1982). *Judgment under uncertainty: Heuristics and biases*. *Science*. Vol. 185. New York: Cambridge University Press. pp. 1124–1131. Bibcode:1974Sci...185.1124T. doi:10.1126/science.185.4157.1124. ISBN 978-0-521-28414-1. PMID 17835457. S2CID 143452957.
- Katz, Yarden (1 November 2012). "Noam Chomsky on Where Artificial Intelligence Went Wrong". *The Atlantic*. Archived from the original on 28 February 2019. Retrieved 26 October 2014.
- Kurzweil, Ray (2005). *The Singularity is Near*. Penguin Books. ISBN 978-0-670-03384-3.
- Langley, Pat (2011). "The changing science of machine learning". *Machine Learning*. **82** (3): 275–279. doi:10.1007/s10994-011-5242-y.
- Legg, Shane; Hutter, Marcus (15 June 2007). "A Collection of Definitions of Intelligence". arXiv:0706.3639 [cs.AI].
- Lenat, Douglas; Guha, R. V. (1989). *Building Large Knowledge-Based Systems*. Addison-Wesley. ISBN 978-0-201-51752-1.
- Lighthill, James (1973). "Artificial Intelligence: A General Survey". *Artificial Intelligence: a paper symposium*. Science Research Council.
- Lombardo, P; Boehm, I; Nairz, K (2020). "RadioComics – Santa Claus and the future of radiology". *Eur J Radiol*. **122** (1): 108771. doi:10.1016/j.ejrad.2019.108771. PMID 31835078.
- Lungarella, M.; Metta, G.; Pfeifer, R.; Sandini, G. (2003). "Developmental robotics: a survey". *Connection Science*. **15** (4): 151–190. CiteSeerX 10.1.1.83.7615. doi:10.1080/09540090310001655110. S2CID 1452734.
- Maker, Meg Houston (2006). "AI@50: AI Past, Present, Future". Dartmouth College. Archived from the original on 3 January 2007. Retrieved 16 October 2008.
- McCarthy, John; Minsky, Marvin; Rochester, Nathan; Shannon, Claude (1955). "A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence". Archived from the original on 26 August 2007. Retrieved 30 August 2007.
- Minsky, Marvin (1967). *Computation: Finite and Infinite Machines*. Englewood Cliffs, N.J.: Prentice-Hall. ISBN 978-0-13-165449-5. Archived from the original on 26 July 2020. Retrieved 18 November 2019.
- Moravec, Hans (1988). *Mind Children*. Harvard University Press. ISBN 978-0-674-57616-2. Archived from the original on 26 July 2020. Retrieved 18 November 2019.
- NRC (United States National Research Council) (1999). "Developments in Artificial Intelligence". *Funding a Revolution: Government Support for Computing Research*. National Academy Press.
- Newell, Allen; Simon, H. A. (1976). "Computer Science as Empirical Inquiry: Symbols and Search". *Communications of the ACM*. **19** (3): 113–126. doi:10.1145/360018.360022..
- Nilsson, Nils (1983). "Artificial Intelligence Prepares for 2001" (PDF). *AI Magazine*. **1** (1). Archived (PDF) from the original on 17 August 2020. Retrieved 22 August 2020. Presidential Address to the Association for the Advancement of Artificial Intelligence.
- Oudeyer, P-Y. (2010). "On the impact of robotics in behavioral and cognitive sciences: from insect navigation to human cognitive development" (PDF). *IEEE Transactions on Autonomous Mental Development*. **2** (1): 2–16. doi:10.1109/tamd.2009.2039057. S2CID 6362217. Archived (PDF) from the original on 3 October 2018. Retrieved 4 June 2013.
- Schank, Roger C. (1991). "Where's the AI". *AI magazine*. Vol. 12, no. 4.
- Searle, John (1980). "Minds, Brains and Programs" (PDF). *Behavioral and Brain Sciences*. **3** (3): 417–457. doi:10.1017/S0140525X00005756. S2CID 55303721. Archived (PDF) from the original on 17 March 2019. Retrieved 22 August 2020.
- Searle, John (1999). *Mind, language and society*. New York: Basic Books. ISBN 978-0-465-04521-1. OCLC 231867665. Archived from the original on 26 July 2020. Retrieved 22 August 2020.
- Simon, H. A. (1965). *The Shape of Automation for Men and Management*. New York: Harper & Row. Archived from the original on 26 July 2020. Retrieved 18 November 2019.
- Solomonoff, Ray (1956). *An Inductive Inference Machine* (PDF). Dartmouth Summer Research Conference on Artificial Intelligence. Archived (PDF) from the original on 26 April 2011. Retrieved 22 March 2011 – via std.com, pdf scanned copy of the original. Later published as
  Solomonoff, Ray (1957). "An Inductive Inference Machine". *IRE Convention Record*. Vol. Section on Information Theory, part 2. pp. 56–62.
- Spadafora, Anthony (21 October 2016). "Stephen Hawking believes AI could be mankind's last accomplishment". *BetaNews*. Archived from the original on 28 August 2017.
- Tao, Jianhua; Tan, Tieniu (2005). *Affective Computing and Intelligent Interaction*. Affective Computing: A Review. Lecture Notes in Computer Science. Vol. LNCS 3784. Springer. pp. 981–995. doi:10.1007/11573548. ISBN 978-3-540-29621-8.
- Tecuci, Gheorghe (March–April 2012). "Artificial Intelligence". *Wiley Interdisciplinary Reviews: Computational Statistics*. **4** (2): 168–180. doi:10.1002/wics.200. S2CID 196141190.
- Thro, Ellen (1993). *Robotics: The Marriage of Computers and Machines*. New York: Facts on File. ISBN 978-0-8160-2628-9. Archived from the original on 26 July 2020. Retrieved 22 August 2020.
- Turing, Alan (October 1950), "Computing Machinery and Intelligence", *Mind*, **LIX** (236): 433–460, doi:10.1093/mind/LIX.236.433, ISSN 0026-4423.
- *UNESCO Science Report: the Race Against Time for Smarter Development*. Paris: UNESCO. 11 June 2021. ISBN 978-92-3-100450-6. Archived from the original on 18 June 2022. Retrieved 18 September 2021.
- Vinge, Vernor (1993). "The Coming Technological Singularity: How to Survive in the Post-Human Era". *Vision 21: Interdisciplinary Science and Engineering in the Era of Cyberspace*: 11. Bibcode:1993vise.nasa...11V. Archived from the original on 1 January 2007. Retrieved 14 November 2011.
- Wason, P. C.; Shapiro, D. (1966). "Reasoning". In Foss, B. M. (ed.). *New horizons in psychology*. Harmondsworth: Penguin. Archived from the original on 26 July 2020. Retrieved 18 November 2019.
- Weng, J.; McClelland; Pentland, A.; Sporns, O.; Stockman, I.; Sur, M.; Thelen, E. (2001). "Autonomous mental development by robots and animals" (PDF). *Science*. **291** (5504): 599–600. doi:10.1126/science.291.5504.599. PMID 11229402. S2CID 54131797. Archived (PDF) from the original on 4 September 2013. Retrieved 4 June 2013 – via msu.edu.

Further reading

- Autor, David H., "Why Are There Still So Many Jobs? The History and Future of Workplace Automation" (2015) 29(3) *Journal of Economic Perspectives* 3.
- Boden, Margaret, *Mind As Machine*, Oxford University Press, 2006.
- Cukier, Kenneth, "Ready for Robots? How to Think about the Future of AI", *Foreign Affairs*, vol. 98, no. 4 (July/August 2019), pp. 192–98. George Dyson, historian of computing, writes (in what might be called "Dyson's Law") that "Any system simple enough to be understandable will not be complicated enough to behave intelligently, while any system complicated enough to behave intelligently will be too complicated to understand." (p. 197.) Computer scientist Alex Pentland writes: "Current AI machine-learning algorithms are, at their core, dead simple stupid. They work, but they work by brute force." (p. 198.)
- Domingos, Pedro, "Our Digital Doubles: AI will serve our species, not control it", *Scientific American*, vol. 319, no. 3 (September 2018), pp. 88–93.
- Gopnik, Alison, "Making AI More Human: Artificial intelligence has staged a revival by starting to incorporate what we know about how children learn", *Scientific American*, vol. 316, no. 6 (June 2017), pp. 60–65.
- Halpern, Sue, "The Human Costs of AI" (review of Kate Crawford, *Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence*, Yale University Press, 2021, 327 pp.; Simon Chesterman, *We, the Robots?: Regulating Artificial Intelligence and the Limits of the Law*, Cambridge University Press, 2021, 289 pp.; Keven Roose, *Futureproof: 9 Rules for Humans in the Age of Automation*, Random House, 217 pp.; Erik J. Larson, *The Myth of Artificial Intelligence: Why Computers Can't Think the Way We Do*, Belknap Press / Harvard University Press, 312 pp.), *The New York Review of Books*, vol. LXVIII, no. 16 (21 October 2021), pp. 29–31. "AI training models can replicate entrenched social and cultural biases. [...] Machines only know what they know from the data they have been given. [p. 30.] [A]rtificial general intelligence–machine-based intelligence that matches our own–is beyond the capacity of algorithmic machine learning... 'Your brain is one piece in a broader system which includes your body, your environment, other humans, and culture as a whole.' [E]ven machines that master the tasks they are trained to perform can't jump domains. AIVA, for example, can't drive a car even though it can write music (and wouldn't even be able to do that without Bach and Beethoven [and other composers on which AIVA is trained])." (p. 31.)
- Johnston, John (2008) *The Allure of Machinic Life: Cybernetics, Artificial Life, and the New AI*, MIT Press.
- Koch, Christof, "Proust among the Machines", *Scientific American*, vol. 321, no. 6 (December 2019), pp. 46–49. Christof Koch doubts the possibility of "intelligent" machines attaining consciousness, because "[e]ven the most sophisticated brain simulations are unlikely to produce conscious feelings." (p. 48.) According to Koch, "Whether machines can become sentient [is important] for ethical reasons. If computers experience life through their own senses, they cease to be purely a means to an end determined by their usefulness to... humans. Per GNW [the Global Neuronal Workspace theory], they turn from mere objects into subjects... with a point of view.... Once computers' cognitive abilities rival those of humanity, their impulse to push for legal and political rights will become irresistible—the right not to be deleted, not to have their memories wiped clean, not to suffer pain and degradation. The alternative, embodied by IIT [Integrated Information Theory], is that computers will remain only supersophisticated machinery, ghostlike empty shells, devoid of what we value most: the feeling of life itself." (p. 49.)
- Marcus, Gary, "Am I Human?: Researchers need new ways to distinguish artificial intelligence from the natural kind", *Scientific American*, vol. 316, no. 3 (March 2017), pp. 58–63. A stumbling block to AI has been an incapacity for reliable disambiguation. An example is the "pronoun disambiguation problem": a machine has no way of determining to whom or what a pronoun in a sentence refers. (p. 61.)
- Gary Marcus, "Artificial Confidence: Even the newest, buzziest systems of artificial general intelligence are stymmied by the same old problems", *Scientific American*, vol. 327, no. 4 (October 2022), pp. 42–45.
- E McGaughey, 'Will Robots Automate Your Job Away? Full Employment, Basic Income, and Economic Democracy' (2018) SSRN, part 2(3) Archived 24 May 2018 at the Wayback Machine.
- George Musser, "Artificial Imagination: How machines could learn creativity and common sense, among other human qualities", *Scientific American*, vol. 320, no. 5 (May 2019), pp. 58–63.
- Myers, Courtney Boyd ed. (2009). "The AI Report" Archived 29 July 2017 at the Wayback Machine. *Forbes* June 2009
- Raphael, Bertram (1976). *The Thinking Computer*. W.H. Freeman and Co. ISBN 978-0716707233. Archived from the original on 26 July 2020. Retrieved 22 August 2020.
- Scharre, Paul, "Killer Apps: The Real Dangers of an AI Arms Race", *Foreign Affairs*, vol. 98, no. 3 (May/June 2019), pp. 135–44. "Today's AI technologies are powerful but unreliable. Rules-based systems cannot deal with circumstances their programmers did not anticipate. Learning systems are limited by the data on which they were trained. AI failures have already led to tragedy. Advanced autopilot features in cars, although they perform well in some circumstances, have driven cars without warning into trucks, concrete barriers, and parked cars. In the wrong situation, AI systems go from supersmart to superdumb in an instant. When an enemy is trying to manipulate and hack an AI system, the risks are even greater." (p. 140.)
- Serenko, Alexander (2010). "The development of an AI journal ranking based on the revealed preference approach" (PDF). *Journal of Informetrics*. **4** (4): 447–59. doi:10.1016/j.joi.2010.04.001. Archived (PDF) from the original on 4 October 2013. Retrieved 24 August 2013.
- Serenko, Alexander; Michael Dohan (2011). "Comparing the expert survey and citation impact journal ranking methods: Example from the field of Artificial Intelligence" (PDF). *Journal of Informetrics*. **5** (4): 629–49. doi:10.1016/j.joi.2011.06.002. Archived (PDF) from the original on 4 October 2013. Retrieved 12 September 2013.
- Tom Simonite (29 December 2014). "2014 in Computing: Breakthroughs in Artificial Intelligence". *MIT Technology Review*. Archived from the original on 2 January 2015.
- Sun, R. & Bookman, L. (eds.), *Computational Architectures: Integrating Neural and Symbolic Processes*. Kluwer Academic Publishers, Needham, MA. 1994.
- Taylor, Paul, "Insanely Complicated, Hopelessly Inadequate" (review of Brian Cantwell Smith, *The Promise of Artificial Intelligence: Reckoning and Judgment*, MIT, 2019, ISBN 978-0262043045, 157 pp.; Gary Marcus and Ernest Davis, *Rebooting AI: Building Artificial Intelligence We Can Trust*, Ballantine, 2019, ISBN 978-1524748258, 304 pp.; Judea Pearl and Dana Mackenzie, *The Book of Why: The New Science of Cause and Effect*, Penguin, 2019, ISBN 978-0141982410, 418 pp.), *London Review of Books*, vol. 43, no. 2 (21 January 2021), pp. 37–39. Paul Taylor writes (p. 39): "Perhaps there is a limit to what a computer can do without knowing that it is manipulating imperfect representations of an external reality."
- Tooze, Adam, "Democracy and Its Discontents", *The New York Review of Books*, vol. LXVI, no. 10 (6 June 2019), pp. 52–53, 56–57. "Democracy has no clear answer for the mindless operation of bureaucratic and technological power. We may indeed be witnessing its extension in the form of artificial intelligence and robotics. Likewise, after decades of dire warning, the environmental problem remains fundamentally unaddressed.... Bureaucratic overreach and environmental catastrophe are precisely the kinds of slow-moving existential challenges that democracies deal with very badly.... Finally, there is the threat du jour: corporations and the technologies they promote." (pp. 56–57.)

External links

Artificial intelligence
at Wikipedia's sister projects

- Definitions from Wiktionary
- Media from Commons
- Quotations from Wikiquote

Textbooks from Wikibooks

Resources from Wikiversity

Data from Wikidata

- "Artificial Intelligence". *Internet Encyclopedia of Philosophy*.
- Thomason, Richmond. "Logic and Artificial Intelligence". In Zalta, Edward N. (ed.). *Stanford Encyclopedia of Philosophy*.
- Artificial Intelligence. BBC Radio 4 discussion with John Agar, Alison Adam & Igor Aleksander (*In Our Time*, 8 December 2005).

Articles related to Artificial intelligence

v · t · e	John McCarthy
	Artificial intelligence · Circumscription · Dartmouth workshop · Frame problem · Garbage collection · Lisp · ALGOL 60 · McCarthy Formalism · McCarthy 91 function · Situation calculus · Space fountain

v · t · e	Philosophy of mind
Philosophers	Anscombe · Austin · Aquinas · Bain · Bergson · Bhattacharya · Block · Brentano · Broad · Burge · Chalmers · Churchland · Dennett · Dharmakirti · Davidson · Descartes · Goldman · Heidegger · Husserl · Feyerabend · Fodor · James · Kierkegaard · Leibniz · Lewis · McDowell · Merleau-Ponty · Minsky · Moore · Nagel · Parfit · Putnam · Popper · Rorty · Ryle · Searle · Spinoza · Turing · Vasubandhu · Wittgenstein · Zhuangzi · *more...*
Theories	Behaviorism · Biological naturalism · Dualism · Eliminative materialism · Emergent materialism · Epiphenomenalism · Functionalism · Idealism · Interactionism · Materialism · Monism · Naïve realism (Neurophenomenology) · Neutral monism · Occasionalism · Parallelism · Phenomenalism · Phenomenology · Physicalism (identity theory) · Property dualism · Representational · Solipsism · Substance dualism
Concepts	Abstract object · **Artificial intelligence** · Chinese room · Cognition · Cognitive closure · Concept · Concept and object · Consciousness · Hard problem of consciousness · Hypostatic abstraction · Idea · Identity · Ingenuity · Intelligence · Intentionality · Introspection · Intuition · Language of thought · Materialism · Mental event · Mental image · Mental process · Mental property · Mental representation · Mind · Mind–body problem · New mysterianism · Pain · Problem of other minds · Propositional attitude · Qualia · Tabula rasa · Understanding · Zombie · *more...*
Related	Metaphysics · Philosophy of artificial intelligence / information / perception / self
	Category · Philosophers category · Project · Task Force

v · t · e	Philosophy of science
Concepts	Analysis · Analytic–synthetic distinction · *A priori* and *a posteriori* · Causality · Commensurability · Consilience · Construct · Creative synthesis · Demarcation problem · Empirical evidence · Explanatory power · Fact · Falsifiability · Feminist method · Functional contextualism · *Ignoramus et ignorabimus* · Inductive reasoning · Intertheoretic reduction · Inquiry · Nature · Objectivity · Observation · Paradigm · Problem of induction · Scientific law · Scientific method · Scientific pluralism · Scientific revolution · Scientific theory · Testability · Theory choice · Theory-ladenness · Underdetermination · Unity of science · more...
Theories	Coherentism · Confirmation holism · Constructive empiricism · Constructive realism · Constructivist epistemology · Contextualism · Conventionalism · Deductive-nomological model · Hypothetico-deductive model · Inductionism · Epistemological anarchism · Evolutionism · Fallibilism · Foundationalism · Instrumentalism · Pragmatism · Model-dependent realism · Naturalism · Physicalism · Positivism / Reductionism / Determinism · Rationalism / Empiricism · Received view / Semantic view of theories · Scientific realism / Anti-realism · Scientific essentialism · Scientific formalism · Scientific skepticism · Scientism · Structuralism · Uniformitarianism · Vitalism
Philosophy of...	Biology · Chemistry · Physics · Social science (Archaeology · Economics · Geography · History · Linguistics · Psychology)
Related topics	Criticism of science · Descriptive science · Faith and rationality · Hard and soft science · History and philosophy of science · Normative science · Pseudoscience · Relationship between religion and science · Rhetoric of science · Science studies · Sociology of scientific ignorance · Sociology of scientific knowledge
Philosophers of science	**Precursors**: Roger Bacon · Francis Bacon · Galileo Galilei · Isaac Newton Auguste Comte · Henri Poincaré · Pierre Duhem · Rudolf Steiner · Karl Pearson · Charles Sanders Peirce · Wilhelm Windelband · Alfred North Whitehead · Bertrand Russell · Otto Neurath · C. D. Broad · Michael Polanyi · Hans Reichenbach · Rudolf Carnap · Karl Popper · Carl Gustav Hempel · W. V. O. Quine · Thomas Kuhn · Imre Lakatos · Paul Feyerabend · Ian Hacking · Bas van Fraassen · Larry Laudan
	Category · Philosophy portal · Science portal

v · t · e	Evolutionary computation
Main Topics	Convergence (evolutionary computing) · Evolutionary algorithm · Evolutionary data mining · Evolutionary multimodal optimization · Human-based evolutionary computation · Interactive evolutionary computation
Algorithms	Cellular evolutionary algorithm · Covariance Matrix Adaptation Evolution Strategy (CMA-ES) · Differential evolution · Evolutionary programming · Genetic algorithm · Genetic programming · Gene expression programming · Evolution strategy · Natural evolution strategy · Neuroevolution · Learning classifier system
Related techniques	Swarm intelligence · Ant colony optimization · Bees algorithm · Cuckoo search · Particle swarm optimization · Bacterial Colony Optimization
Metaheuristic methods	Firefly algorithm · Harmony search · Gaussian adaptation · Memetic algorithm
Related topics	Artificial development · **Artificial intelligence** · Artificial life · Digital organism · Evolutionary robotics · Fitness function · Fitness landscape · Fitness approximation · Genetic operators · Interactive evolutionary computation · No free lunch in search and optimization · Machine learning · Mating pool · Program synthesis
Journals	Evolutionary Computation (journal)

v · t · e	Differentiable computing
General	**Differentiable programming** · Information geometry · Statistical manifold Automatic differentiation · Neuromorphic engineering · Pattern recognition · Tensor calculus · Computational learning theory · Inductive bias
Concepts	Gradient descent (SGD) · Clustering · Regression (Overfitting) · Hallucination · Adversary · Attention · Convolution · Loss functions · Backpropagation · Normalization · Activation (Softmax · Sigmoid · Rectifier) · Regularization · Datasets (Augmentation) · Diffusion · Autoregression
Applications	Machine learning (In-context learning) · Artificial neural network (Deep learning) · Scientific computing · **Artificial Intelligence** · Language model (Large language model)
Hardware	IPU · TPU · VPU · Memristor · SpiNNaker
Software libraries	TensorFlow · PyTorch · Keras · Theano · JAX · LangChain
Implementations	**Audio–visual**: AlexNet · WaveNet · Human image synthesis · HWR · OCR · Speech synthesis · Speech recognition · Facial recognition · AlphaFold · DALL-E · Midjourney · Stable Diffusion **Verbal**: Word2vec · Seq2seq · BERT · LaMDA (Bard) · NMT · Project Debater · IBM Watson · GPT family (GPT-2 · GPT-3 · ChatGPT · GPT-4) · GPT-J · Chinchilla AI · PaLM · BLOOM · LLaMA **Decisional**: AlphaGo · AlphaZero · Q-learning · SARSA · OpenAI Five · Self-driving car · MuZero · Action selection (Auto-GPT) · Robot control
People	Yoshua Bengio · Alex Graves · Ian Goodfellow · Stephen Grossberg · Demis Hassabis · Geoffrey Hinton · Yann LeCun · Fei-Fei Li · Andrew Ng · Jürgen Schmidhuber · David Silver
Organizations	Anthropic · EleutherAI · Google DeepMind · OpenAI · Meta AI · Mila · MIT CSAIL
Architectures	Neural Turing machine · Differentiable neural computer · Transformer · Recurrent neural network (RNN) · Long short-term memory (LSTM) · Gated recurrent unit (GRU) · Echo state network · Multilayer perceptron (MLP) · Convolutional neural network · Residual network · Autoencoder · Variational autoencoder (VAE) · Generative adversarial network (GAN) · Graph neural network
	Portals (Computer programming · Technology) · Categories (Artificial neural networks · Machine learning)

v · t · e	Computer science
	Note: This template roughly follows the 2012 ACM Computing Classification System.
Hardware	Printed circuit board · Peripheral · Integrated circuit · Very Large Scale Integration · Systems on Chip (SoCs) · Energy consumption (Green computing) · Electronic design automation · Hardware acceleration
Computer systems organization	Computer architecture · Embedded system · Real-time computing · Dependability
Networks	Network architecture · Network protocol · Network components · Network scheduler · Network performance evaluation · Network service
Software organization	Interpreter · Middleware · Virtual machine · Operating system · Software quality
Software notations and tools	Programming paradigm · Programming language · Compiler · Domain-specific language · Modeling language · Software framework · Integrated development environment · Software configuration management · Software library · Software repository
Software development	Control variable · Software development process · Requirements analysis · Software design · Software construction · Software deployment · Software engineering · Software maintenance · Programming team · Open-source model
Theory of computation	Model of computation · Formal language · Automata theory · Computability theory · Computational complexity theory · Logic · Semantics
Algorithms	Algorithm design · Analysis of algorithms · Algorithmic efficiency · Randomized algorithm · Computational geometry
Mathematics of computing	Discrete mathematics · Probability · Statistics · Mathematical software · Information theory · Mathematical analysis · Numerical analysis · Theoretical computer science
Information systems	Database management system · Information storage systems · Enterprise information system · Social information systems · Geographic information system · Decision support system · Process control system · Multimedia information system · Data mining · Digital library · Computing platform · Digital marketing · World Wide Web · Information retrieval
Security	Cryptography · Formal methods · Security services · Intrusion detection system · Hardware security · Network security · Information security · Application security
Human–computer interaction	Interaction design · Social computing · Ubiquitous computing · Visualization · Accessibility · Synthography
Concurrency	Concurrent computing · Parallel computing · Distributed computing · Multithreading · Multiprocessing
Artificial intelligence	Natural language processing · Knowledge representation and reasoning · Computer vision · Automated planning and scheduling · Search methodology · Control method · Philosophy of artificial intelligence · Distributed artificial intelligence
Machine learning	Supervised learning · Unsupervised learning · Reinforcement learning · Multi-task learning · Cross-validation
Graphics	Animation · Rendering · Photograph manipulation · Graphics processing unit · Mixed reality · Virtual reality · Image compression · Solid modeling
Applied computing	E-commerce · Enterprise software · Computational mathematics · Computational physics · Computational chemistry · Computational biology · Computational social science · Computational engineering · Computational healthcare · Digital art · Electronic publishing · Cyberwarfare · Electronic voting · Video games · Word processing · Operations research · Educational technology · Document management
	Category · Outline · WikiProject · Commons

v · t · e

Emerging technologies

Fields	**Information and communications**	Ambient intelligence (Internet of things) · **Artificial intelligence** (Applications of artificial intelligence · Progress in artificial intelligence · Machine translation · Mobile translation · Machine vision · Semantic Web · Speech recognition) · Atomtronics · Carbon nanotube field-effect transistor · Cybermethodology · Fourth-generation optical discs (3D optical data storage · Holographic data storage) · GPGPU · Memory (CBRAM · FRAM · Millipede · MRAM · NRAM · PRAM · Racetrack memory · RRAM · SONOS · ECRAM · UltraRAM) · Optical computing · RFID (Chipless RFID) · Software-defined radio · Three-dimensional integrated circuit
Topics		Collingridge dilemma · Differential technological development · Disruptive innovation · Ephemeralization · Ethics (Bioethics · Cyberethics · Neuroethics · Robot ethics) · Exploratory engineering · Fictional technology · Proactionary principle · Technological change (Technological unemployment) · Technological convergence · Technological evolution · Technological paradigm · Technology forecasting (Accelerating change · Future-oriented technology analysis · Horizon scanning · Moore's law · Technological singularity · Technology scouting) · Technology readiness level · Technology roadmap · Transhumanism

Category · List

v · t · e

Robotics

Main articles	Outline · Glossary · Index · History · Geography · Hall of Fame · Ethics · Laws · Competitions · AI competitions
Types	Aerobot · Anthropomorphic (Humanoid · Android · Cyborg · Gynoid) · Claytronics · Companion · Animatronic (Audio-Animatronics) · Industrial · Articulated (arm) · Domestic · Educational · Entertainment · Juggling · Military · Medical · Service · Disability · Agricultural · Food service · Retail · BEAM robotics · Soft robotics
Classifications	Biorobotics · Unmanned vehicle (aerial · ground) · Mobile robot · Microbotics · Nanorobotics · Necrobotics · Robotic spacecraft (Space probe) · Swarm · Telerobotics · Underwater (remotely-operated)
Locomotion	Tracks · Walking (Hexapod) · Climbing · Electric unicycle · Robot navigation
Navigation and mapping	Motion planning · Simultaneous localization and mapping · Visual odometry
Research	Evolutionary · Kits · Simulator · Suite · Open-source · Software · Adaptable · Developmental · Paradigms · Ubiquitous
Related	Critique of work · Powered exoskeleton · Technological unemployment · Terrainability · Fictional robots

Category · Outline

v · t · e

Existential risk from artificial intelligence

Concepts	AI alignment · AI capability control · AI safety · AI takeover · Accelerating change · Existential risk from artificial general intelligence · Friendly artificial intelligence · Instrumental convergence · Intelligence explosion · Machine ethics · Superintelligence · Technological singularity
Organizations	Allen Institute for AI · Alignment Research Center · Center for Applied Rationality · Center for Human-Compatible Artificial Intelligence · Centre for the Study of Existential Risk · Foundational Questions Institute · Future of Humanity Institute · Future of Life Institute · Google DeepMind · Humanity+ · Institute for Ethics and Emerging Technologies · Leverhulme Centre for the Future of Intelligence · Machine Intelligence Research Institute · OpenAI
People	Scott Alexander · Nick Bostrom · Eric Drexler · Sam Harris · Stephen Hawking · Bill Hibbard · Bill Joy · Elon Musk · Steve Omohundro · Huw Price · Martin Rees · Stuart J. Russell · Jaan Tallinn · Max Tegmark · Frank Wilczek · Roman Yampolskiy · Andrew Yang · Eliezer Yudkowsky
Other	Artificial intelligence as a global catastrophic risk · Controversies and dangers of artificial general intelligence · Ethics of artificial intelligence · Suffering risks · *Human Compatible* · Open Letter on Artificial Intelligence · *Our Final Invention* · *The Precipice* · *Superintelligence: Paths, Dangers, Strategies* · *Do You Trust This Computer?* · Artificial Intelligence Act

Category

v · t · e

Subfields of and cyberneticians involved in cybernetics

Subfields	**Artificial intelligence** · Biological cybernetics · Biomedical cybernetics · Biorobotics · Biosemiotics · Neurocybernetics · Catastrophe theory · Computational neuroscience · Connectionism · Control theory · Cybernetics in the Soviet Union · Decision theory · Emergence · Engineering cybernetics · Homeostasis · Information theory · Management cybernetics · Medical cybernetics · Second-order cybernetics · Semiotics · Sociocybernetics · Polycontexturality · Synergetics
Cyberneticians	Alexander Lerner · Alexey Lyapunov · Alfred Radcliffe-Brown · Allenna Leonard · Anthony Wilden · Buckminster Fuller · Charles François · Genevieve Bell · Margaret Boden · Claude Bernard · Cliff Joslyn · Erich von Holst · Ernst von Glasersfeld · Francis Heylighen · Francisco Varela · Frederic Vester · Charles Geoffrey Vickers · Gordon Pask · Gordon S. Brown · Gregory Bateson · Heinz von Foerster · Humberto Maturana · I. A. Richards · Igor Aleksander · Jacque Fresco · Jakob von Uexküll · Jason Jixuan Hu · Jay Wright Forrester · Jennifer Wilby · John N. Warfield · Kevin Warwick · Ludwig von Bertalanffy · Maleyka Abbaszadeh · Manfred Clynes · Margaret Mead · Marian Mazur · N. Katherine Hayles · Natalia Bekhtereva · Niklas Luhmann · Norbert Wiener · Pyotr Grigorenko · Qian Xuesen · Ranulph Glanville · Robert Trappl · Sergei P. Kurdyumov · Anthony Stafford Beer · Stuart Kauffman · Stuart Umpleby · Talcott Parsons · Ulla Mitzdorf · Valentin Turchin · Valentin Braitenberg · William Ross Ashby · Walter Bradford Cannon · Walter Pitts · Warren McCulloch · William Grey Walter

v · t · e

Glossaries of science and engineering

Aerospace engineering · Agriculture · Archaeology · Architecture · Artificial intelligence · Astronomy · Biology · Botany · Calculus · Cell biology · Chemistry · Civil engineering · Clinical research · Computer hardware · Computer science · Developmental and reproductive biology · Ecology · Economics · Electrical and electronics engineering · Engineering (A–L · M–Z) · Entomology · Environmental science · Evolutionary biology · Genetics (0–L · M–Z) · Geography (Arabic toponyms · Hebrew toponyms · Western and South Asia) · Geology · Ichthyology · Machine vision · Mathematics · Mechanical engineering · Medicine · Meteorology · Mycology · Nanotechnology · Ornithology · Physics · Probability and statistics · Psychiatry · Quantum computing · Robotics · Scientific naming · Structural engineering · Virology

Authority control: National Spain · France · BnF data · Germany · Israel · United States · Latvia · Japan · Czech Republic

Categories: Artificial intelligence | Cybernetics | Computational neuroscience | Computational fields of study | Data science | Emerging technologies | Formal sciences | Intelligence by type | Unsolved problems in computer science