Artificial intelligence

Artificial intelligence (**AI**), is intelligence demonstrated by machines, unlike the **natural intelligence** displayed by humans and animals. Leading AI textbooks define the field as the study of "intelligent agents": any device that perceives its environment and takes actions that maximize its chance of successfully achieving its goals.[3] Colloquially, the term "artificial intelligence" is often used to describe machines (or computers) that mimic "cognitive" functions that humans associate with the human mind, such as "learning" and "problem solving".[4]

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.[5] A quip in Tesler's Theorem says "AI is whatever hasn't been done yet."[6] For instance, optical character recognition is frequently excluded from things considered to be AI,[7] having become a routine technology.[8] Modern machine capabilities generally classified as AI include successfully understanding human speech,[9] competing at the highest level in strategic game systems (such as chess and Go),[10] autonomously operating cars, intelligent routing in content delivery networks, and military simulations.[11]

Artificial intelligence was founded as an academic discipline in 1955, and in the years since has experienced several waves of optimism,[12][13] followed by disappointment and the loss of funding (known as an "AI winter"),[14][15] followed by new approaches, success and renewed funding.[13][16] For most of its history, AI research has been divided into sub-fields that often fail to communicate with each other.[17] These sub-fields are based on technical considerations, such as particular goals (e.g. "robotics" or "machine learning"),[18] the use of particular tools ("logic" or artificial neural networks), or deep philosophical differences.[21][22][23] Sub-fields have also been based on social factors (particular institutions or the work of particular researchers).[17]

The traditional problems (or goals) of AI research include reasoning, knowledge representation, planning, learning, natural language processing, perception and the ability to move and manipulate objects.[18] General intelligence is among the field's long-term goals.[24] Approaches include statistical methods, computational intelligence, and traditional symbolic AI. Many tools are used in AI, including versions of search and mathematical optimization, artificial neural networks, and methods based on statistics, probability and economics. The AI field draws upon computer science, information engineering, mathematics, 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".[25] This raises philosophical arguments about the mind and the ethics of creating artificial beings endowed with human-like intelligence. These issues have been explored by myth, fiction and philosophy since antiquity.[30] Some people also consider AI to be a danger to humanity if it progresses unabated.[31][32] Others believe that AI, unlike previous technological revolutions, will create a risk of mass unemployment.[33]

In the twenty-first century, AI techniques have experienced a resurgence following concurrent advances in computer power, large amounts of data, and theoretical understanding; and AI techniques have become an essential part of the technology industry, helping to solve many challenging problems in computer science, software engineering and operations research.[34][16]

Contents

History

Thought-capable artificial beings appeared as storytelling devices in antiquity,[35] and have been common in fiction, as in Mary Shelley's *Frankenstein* or Karel Čapek's *R.U.R. (Rossum's Universal Robots)*.[36] These characters and their fates raised many of the same issues now discussed in the ethics of artificial intelligence.[30]

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

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.[37] Along with concurrent discoveries in neurobiology, information theory and cybernetics, this led researchers to consider the possibility of building an electronic brain. Turing proposed changing the question from whether a machine was intelligent, to "whether or not it is possible for machinery to show intelligent behaviour".[38] The first work that is now generally recognized as AI was McCullouch and Pitts' 1943 formal design for Turing-complete "artificial neurons".[39]

The field of AI research was born at a workshop at Dartmouth College in 1956,[40] where the term "Artificial Intelligence" was coined by John McCarthy to distinguish the field from cybernetics and escape the influence of the cyberneticist Norbert Wiener.[41] Attendees Allen Newell (CMU), Herbert Simon (CMU), John McCarthy (MIT), Marvin Minsky (MIT) and Arthur Samuel (IBM) became the founders and leaders of AI research.[42] They and their students produced programs that the press described as "astonishing":[43] computers were learning checkers strategies (c. 1954)[44] (and by 1959 were reportedly playing better than the average human),[45] solving word problems in algebra, proving logical theorems (Logic Theorist, first run c. 1956) and speaking English.[46] By the middle of the 1960s, research in the U.S. was heavily funded by the Department of Defense[47] and laboratories had been established around the world.[48] AI's founders were optimistic about the future: Herbert Simon predicted, "machines will be capable, within twenty years, of doing any work a man can do". Marvin Minsky agreed, writing, "within a generation ... the problem of creating 'artificial intelligence' will substantially be solved".[12]

They 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[49] 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",[14] a period when obtaining funding for AI projects was difficult.

In the early 1980s, AI research was revived by the commercial success of expert systems,[50] 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.[13] However, beginning with the collapse of the Lisp Machine market in 1987, AI once again fell into disrepute, and a second, longer-lasting hiatus began.[15]

The development of metal–oxide–semiconductor (MOS) very-large-scale integration (VLSI), in the form of complementary MOS (CMOS) transistor technology, enabled the development of practical artificial neural network (ANN) technology in the 1980s. A landmark publication in the field was the 1989 book *Analog VLSI Implementation of Neural Systems* by Carver A. Mead and Mohammed Ismail.[51]

In the late 1990s and early 21st century, AI began to be used for logistics, data mining, medical diagnosis and other areas.[34] The success was due to increasing computational power (see Moore's law and transistor count), greater emphasis on solving specific problems, new ties between AI and other fields (such as statistics, economics and mathematics), and a commitment by researchers to mathematical methods and scientific standards.[52] Deep Blue became the first computer chess-playing system to beat a reigning world chess champion, Garry Kasparov, on 11 May 1997.[53]

In 2011, 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.[54] 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.[55] The Kinect, which provides a 3D body–motion interface for the Xbox 360 and the Xbox One, uses algorithms that emerged from lengthy AI research[56] as do intelligent personal assistants in smartphones.[57] 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.[10][58] In the 2017 Future of Go Summit, AlphaGo won a three-game match with Ke Jie,[59] who at the time continuously held the world No. 1 ranking for two years.[60][61] This marked the completion of a significant milestone in the development of Artificial Intelligence as Go is a relatively complex game, more so than Chess.

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. Clark also presents factual data indicating the improvements of AI since 2012 supported by lower error rates in image processing tasks.[62] He attributes 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.[16] Other cited examples include Microsoft's development of a Skype system that can automatically translate from one language to another and Facebook's system that can describe images to blind people.[62] In a 2017 survey, one in five companies reported they had "incorporated AI in some offerings or processes".[63][64] Around 2016, China greatly accelerated its government funding; given its large supply of data and its rapidly increasing research output, some observers believe it may be on track to becoming an "AI superpower".[65][66] However, it has been acknowledged that reports regarding artificial intelligence have tended to be exaggerated.[67][68][69]

Basics

Computer science defines AI research as the study of "intelligent agents": any device that perceives its environment and takes actions that maximize its chance of successfully achieving its goals.[3] A more elaborate definition characterizes AI as "a system's ability to correctly interpret external data, to learn from such data, and to use those learnings to achieve specific goals and tasks through flexible adaptation."[70]

A typical AI analyzes its environment and takes actions that maximize its chance of success.[3] An AI's intended utility function (or goal) can be simple ("1 if the AI wins a game of Go, 0 otherwise") or complex ("Perform actions mathematically similar to ones that succeeded in the past"). Goals can be explicitly defined or induced. If the AI is programmed for "reinforcement learning", goals can be implicitly induced by rewarding some types of behavior or punishing others.[a] Alternatively, an evolutionary system can induce goals by using a "fitness function" to mutate and preferentially replicate high-scoring AI systems, similar to how animals evolved to innately desire certain goals such as finding food.[71] Some AI systems, such as nearest-neighbor, instead of reason by analogy, these systems are not generally given goals, except to the degree that goals are implicit in their training data.[72] Such systems can still be benchmarked if the non-goal system is framed as a system whose "goal" is to successfully accomplish its narrow classification task.[73]

AI often revolves around the use of algorithms. An algorithm is a set of unambiguous instructions that a mechanical computer can execute.[b] A complex algorithm is often built on top of other, simpler, algorithms. A simple example of an algorithm is the following (optimal for first player) recipe for play at tic-tac-toe:[74]

1. If someone has a "threat" (that is, two in a row), take the remaining square. Otherwise,

2. if a move "forks" to create two threats at once, play that move. Otherwise,
3. take the center square if it is free. Otherwise,
4. if your opponent has played in a corner, take the opposite corner. Otherwise,
5. take an empty corner if one exists. Otherwise,
6. take any empty square.

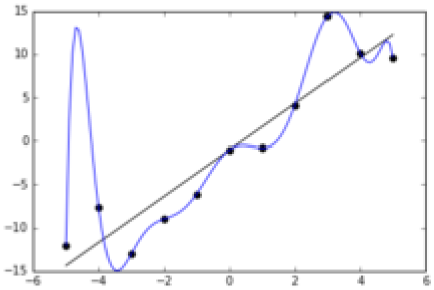
Many AI algorithms are capable of learning from data; they can enhance themselves by learning new heuristics (strategies, or "rules of thumb", that have worked well in the past), or can themselves write other algorithms. Some of the "learners" described below, including Bayesian networks, decision trees, and nearest-neighbor, could theoretically, (given infinite data, time, and memory) learn to approximate any function, including which combination of mathematical functions would best describe the world. These learners could therefore derive all possible knowledge, by considering every possible hypothesis and matching them against the data. In practice, it is seldom possible to consider every possibility, because of the phenomenon of "combinatorial explosion", where the time needed to solve a problem grows exponentially. Much of AI research involves figuring out how to identify and avoid considering a broad range of possibilities unlikely to be beneficial.[75][76] For example, when viewing a map and looking for the shortest driving route from Denver to New York in the East, one can in most cases skip looking at any path through San Francisco or other areas far to the West; thus, an AI wielding a pathfinding algorithm like A* can avoid the combinatorial explosion that would ensue if every possible route had to be ponderously considered.[77]

The earliest (and easiest to understand) approach to AI was symbolism (such as formal logic): "If an otherwise healthy adult has a fever, then they may have influenza". A second, more general, approach is Bayesian inference: "If the current patient has a fever, adjust the probability they have influenza in such-and-such way". The third major approach, extremely popular in routine business AI applications, are analogizers such as SVM and nearest-neighbor: "After examining the records of known past patients whose temperature, symptoms, age, and other factors mostly match the current patient, X% of those patients turned out to have influenza". A fourth approach is harder to intuitively understand, but is inspired by how the brain's machinery works: the artificial neural network approach uses artificial "neurons" that can learn by comparing itself to the desired output and altering the strengths of the connections between its internal neurons to "reinforce" connections that seemed to be useful. These four main approaches can overlap with each other and with evolutionary systems; for example, neural nets can learn to make inferences, to generalize, and to make analogies. Some systems implicitly or explicitly use multiple of these approaches, alongside many other AI and non-AI algorithms; the best approach is often different depending on the problem.[78][79]

Learning algorithms work on the basis that strategies, algorithms, and inferences that worked well in the past are likely to continue working well in the future. These inferences can be obvious, such as "since the sun rose every morning for the last 10,000 days, it will probably rise tomorrow morning as well". They can be nuanced, such as "X% of families have geographically separate species with color variants, so there is a Y% chance that undiscovered black swans exist". Learners also work on the basis of "Occam's razor": The simplest theory that explains the data is the likeliest. Therefore, according to Occam's razor principle, a learner must be designed such that it prefers simpler theories to complex theories, except in cases where the complex theory is proven substantially better.

Settling on a bad, overly complex theory gerrymandered to fit all the past training data is known as overfitting. Many systems attempt to reduce overfitting by rewarding a theory in accordance with how well it fits the data, but penalizing the theory in accordance with how complex the theory is.[80] Besides classic overfitting, learners can also disappoint by "learning the wrong lesson". A toy example is that an image classifier trained only on pictures of brown horses and black cats might conclude that all brown patches are likely to be horses.[81] A real-world example is that, unlike humans, current image classifiers don't determine the spatial relationship between components of the picture; instead, they learn abstract patterns of pixels that humans are oblivious to, but that linearly correlate with images of certain types of real objects. Faintly superimposing such a pattern on a legitimate image results in an "adversarial" image that the system misclassifies.[c][82][83]

Compared with humans, existing AI lacks several features of human "commonsense reasoning"; most notably, humans have powerful mechanisms for reasoning about "naïve physics" such as space, time, and physical interactions. This enables even young children to easily make inferences like "If I roll this pen off a table, it will fall on the floor". Humans also have a powerful mechanism of "folk psychology" that helps them to interpret natural-language sentences such as "The city councilmen refused the demonstrators a permit because they advocated violence" (A generic AI has difficulty discerning whether the ones alleged to be advocating violence are the councilmen or the demonstrators[86][87][88]). This lack of "common knowledge" means that AI often makes different mistakes than humans make, in ways that can seem incomprehensible. For example, existing self-driving cars cannot reason about the location nor the intentions of pedestrians in the exact way that humans do, and instead must use non-human modes of reasoning to avoid accidents.[89][90][91]



The blue line could be an example of overfitting a linear function due to random noise.

A self-driving car system may use a neural network to determine which parts of the picture seem to match previous training images of pedestrians, and then model those areas as slow-moving but somewhat unpredictable rectangular prisms that must be avoided.[84][85]

Challenges

The cognitive capabilities of current architectures are very limited, using only a simplified version of what intelligence is really capable of. For instance, the human mind has come up with ways to reason beyond measure and logical explanations to different occurrences in life. What would have been otherwise straightforward, an equivalently difficult problem may be challenging to solve computationally as opposed to using the human mind. This gives rise to two classes of models: structuralist and functionalist. The structural models aim to loosely mimic the basic intelligence operations of the mind such as reasoning and logic. The functional model refers to the correlating data to its computed counterpart.[92]

The overall research goal of artificial intelligence is to create technology that allows computers and machines to function in an intelligent manner. 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.[18]

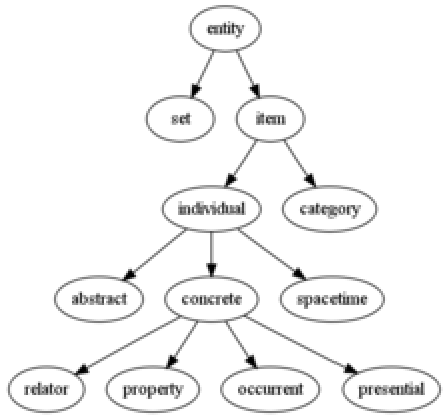
Reasoning, problem solving

Early researchers developed algorithms that imitated step-by-step reasoning that humans use when they solve puzzles or make logical deductions.[93] By the late 1980s and 1990s, AI research had developed methods for dealing with uncertain or incomplete information, employing concepts from probability and economics.[94]

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.[75] 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.[95]

Knowledge representation

Knowledge representation[96] and knowledge engineering[97] are central to classical AI research. Some "expert systems" attempt to gather explicit knowledge possessed by experts in some narrow domain. In addition, some projects attempt to gather the "commonsense knowledge" known to the average person into a database containing extensive knowledge about the world. Among the things a comprehensive commonsense knowledge base would contain are: objects, properties, categories and relations between objects;[98] situations, events, states and time;[99] causes and effects;[100] knowledge about knowledge (what we know about what other people know);[101] and many other, less well researched domains. 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. The semantics of these are captured as description logic concepts, roles, and individuals, and typically implemented as classes, properties, and individuals in the Web Ontology Language.[102] The most general ontologies are called upper ontologies, which attempt to provide a foundation for all other knowledge[103] by acting as mediators between domain ontologies that cover specific knowledge about a particular knowledge domain (field of interest or area of concern). Such formal knowledge representations can be used in content-based indexing and retrieval,[104] scene interpretation,[105] clinical decision support,[106] knowledge discovery (mining "interesting" and actionable inferences from large databases),[107] and other areas.[108]



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

Among the most difficult problems in knowledge representation are:

Default reasoning and the qualification problem

Many of the things people know take the form of "working assumptions". For example, if a bird comes up in conversation, people typically picture a fist-sized animal that sings and flies. None of these things are true about all birds. John McCarthy identified this problem in 1969[109] as the qualification problem: for any commonsense rule that AI researchers care to represent, there tend to be a huge number of exceptions. Almost nothing is simply true or false in the way that abstract logic requires. AI research has explored a number of solutions to this problem.[110]

Breadth of commonsense knowledge

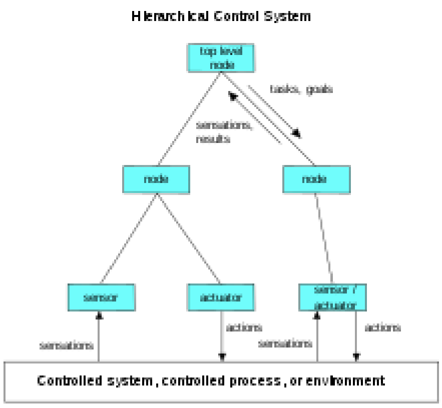
The number of atomic facts that the average person knows is very large. Research projects that attempt to build a complete knowledge base of commonsense knowledge (e.g., Cyc) require enormous amounts of laborious ontological engineering—they must be built, by hand, one complicated concept at a time.[111]

Subsymbolic form of some commonsense knowledge

Much of what people know is not represented as "facts" or "statements" that they could express verbally. For example, a chess master will avoid a particular chess position because it "feels too exposed"[112] or an art critic can take one look at a statue and realize that it is a fake.[113] These are non-conscious and sub-symbolic intuitions or tendencies in the human brain.[114] Knowledge like this informs, supports and provides a context for symbolic, conscious knowledge. As with the related problem of sub-symbolic reasoning, it is hoped that situated AI, computational intelligence, or statistical AI will provide ways to represent this knowledge.[114]

Planning

Intelligent agents must be able to set goals and achieve them.[115] They need a way to visualize the future—a representation of the state of the world and be able to make predictions about how their actions will change it—and be able to make choices that maximize the utility (or "value") of available choices.[116]



A hierarchical control system is a form of control system in which a set of devices and governing software is arranged in a hierarchy.

In classical planning problems, the agent can assume that it is the only system acting in the world, allowing the agent to be certain of the consequences of its actions.[117] However, if the agent is not the only actor, then it requires that the agent can reason under uncertainty. This calls for an agent that can not only assess its environment and make predictions but also evaluate its predictions and adapt based on its assessment.[118]

Multi-agent planning uses the cooperation and competition of many agents to achieve a given goal. Emergent behavior such as this is used by evolutionary algorithms and swarm intelligence.[119]

Learning

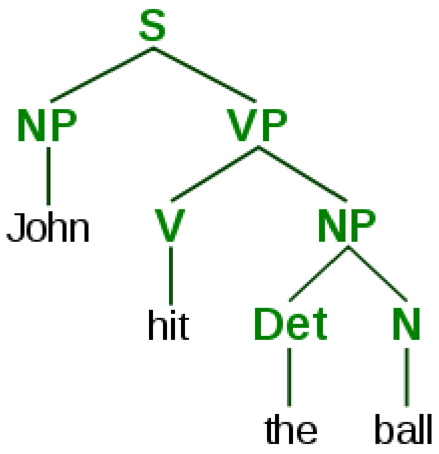

For this project the AI had to find the typical patterns in the colors and brushstrokes of Renaissance painter Raphael. The portrait shows the face of the actress Ornella Muti, "painted" by AI in the style of Raphael.

Machine learning (ML), a fundamental concept of AI research since the field's inception,[122] is the study of computer algorithms that improve automatically through experience.[123][124]

Unsupervised learning is the ability to find patterns in a stream of input, without requiring a human to label the inputs first. Supervised learning includes both classification and numerical regression, which requires a human to label the input data first. Classification is used to determine what category something belongs in, and occurs after a program sees a number of examples of things from several categories. 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.[124] 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". Computational learning theory can assess learners by computational complexity, by sample complexity (how much data is required), or by other notions of optimization.[125] In reinforcement learning[126] the agent is rewarded for good responses and punished for bad ones. The agent uses this sequence of rewards and punishments to form a strategy for operating in its problem space.

Natural language processing

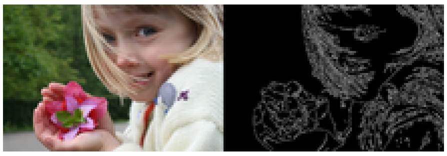
Natural language processing[127] (NLP) allows machines to read and understand human language. A sufficiently powerful natural language processing system would enable natural-language user interfaces and the acquisition of knowledge directly from human-written sources, such as newswire texts. Some straightforward applications of natural language processing include information retrieval, text mining, question answering[128] and machine translation.[129] Many current approaches use word co-occurrence frequencies to construct syntactic representations of text. "Keyword spotting" strategies for search are popular and scalable but dumb; a search query for "dog" might only match documents with the literal word "dog" and miss a document with the word "poodle". "Lexical affinity" strategies use the occurrence of words such as "accident" to assess the sentiment of a document. Modern statistical NLP approaches can combine all these strategies as well as others, and often achieve acceptable accuracy at the page or paragraph level. Beyond semantic NLP, the ultimate goal of "narrative" NLP is to embody a full understanding of commonsense reasoning.[130] By 2019, transformer-based deep learning architectures could generate coherent text.[131]



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

Perception

Machine perception[132] is the ability to use input from sensors (such as cameras (visible spectrum or infrared), microphones, wireless signals, and active lidar, sonar, radar, and tactile sensors) to deduce aspects of the world. Applications include speech recognition,[133] facial recognition, and object recognition.[134] Computer vision is the ability to analyze visual input. Such input is usually ambiguous; a giant, fifty-meter-tall pedestrian far away may produce the same pixels as a nearby normal-sized pedestrian, requiring the AI to judge the relative likelihood and reasonableness of different interpretations, for example by using its "object model" to assess that fifty-meter pedestrians do not exist.[135]



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

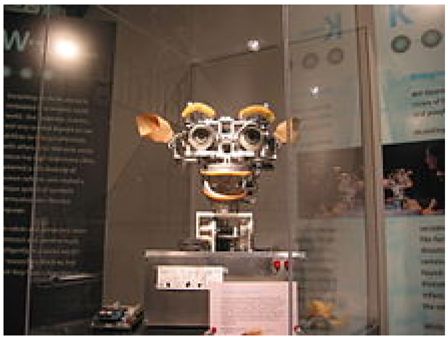
Motion and manipulation

AI is heavily used in robotics.[136] Advanced robotic arms and other industrial robots, widely used in modern factories, can learn from experience how to move efficiently despite the presence of friction and gear slippage.[137] A modern mobile robot, when given a small, static, and visible environment, can easily determine its location and map its environment; however, dynamic environments, such as (in endoscopy) the interior of a patient's breathing body, pose a greater challenge. Motion planning is the process of breaking down a movement task into "primitives" such as individual joint movements. Such movement often involves compliant motion, a process where movement requires maintaining physical contact with an object.[138][139][140] Moravec's paradox generalizes that low-level sensorimotor skills that humans take for granted are, counterintuitively, difficult to program into a robot; the paradox is named after Hans Moravec, who stated in 1988 that "it is comparatively easy to make computers exhibit adult level performance on intelligence tests or playing checkers, and difficult or impossible to give them the skills of a one-year-old when it comes to perception and mobility".[141][142] This is attributed to the fact that, unlike checkers, physical dexterity has been a direct target of natural selection for millions of years.[143]

Social intelligence

Moravec's paradox can be extended to many forms of social intelligence.[145][146] Distributed multi-agent coordination of autonomous vehicles remains a difficult problem.[147] Affective computing is an interdisciplinary umbrella that comprises systems which recognize, interpret, process, or simulate human affects.[148][149][150] Moderate successes related to affective computing include textual sentiment analysis and, more recently, multimodal affect analysis (see multimodal sentiment analysis), wherein AI classifies the affects displayed by a videotaped subject.[151]

In the long run, social skills and an understanding of human emotion and game theory would be valuable to a social agent. The ability to predict the actions of others by understanding their motives and emotional states would allow an agent to make better decisions. Some computer systems mimic human emotion and expressions to appear more sensitive to the emotional dynamics of human interaction, or to otherwise facilitate human–computer interaction.[152] Similarly, some virtual assistants are programmed to speak conversationally or even to banter humorously; this tends to give naïve users an unrealistic conception of how intelligent existing computer agents actually are.[153]



Kismet, a robot with rudimentary social skills[144]

General intelligence

Historically, projects such as the Cyc knowledge base (1984–) and the massive Japanese Fifth Generation Computer Systems initiative (1982–1992) attempted to cover the breadth of human cognition. These early projects failed to escape the limitations of non-quantitative symbolic logic models and, in retrospect, greatly underestimated the difficulty of cross-domain AI. Nowadays, most current AI researchers work instead on tractable "narrow AI" applications (such as medical diagnosis or automobile navigation).[154] Many researchers predict that such "narrow AI" work in different individual domains will eventually be incorporated into a machine with artificial general intelligence (AGI), combining most of the narrow skills mentioned in this article and at some point even exceeding human ability in most or all these areas.[24][155] Many advances have general, cross-domain significance. One high-profile example is that DeepMind in the 2010s developed a "generalized artificial intelligence" that could learn many diverse Atari games on its own, and later developed a variant of the system which succeeds at sequential learning.[156][157][158] Besides transfer learning,[159] hypothetical AGI breakthroughs could include the development of reflective architectures that can engage in decision-theoretic metareasoning, and figuring out how to "slurp up" a comprehensive knowledge base from the entire unstructured Web.[9] Some argue that some kind of (currently-undiscovered) conceptually straightforward, but mathematically difficult, "Master Algorithm" could lead to AGI.[160] Finally, a few "emergent" approaches look to simulating human intelligence extremely closely, and believe that anthropomorphic features like an artificial brain or simulated child development may someday reach a critical point where general intelligence emerges.[161][162]

Many of the problems in this article may also require general intelligence, if machines are to solve the problems as well as people do. For example, even specific straightforward tasks, like machine translation, require that a machine read and write in both languages (NLP), follow the author's argument (reason), know what is being talked about (knowledge), and faithfully reproduce the author's original intent (social intelligence). A problem like machine translation is considered "AI-complete", because all of these problems need to be solved simultaneously in order to reach human-level machine performance.

Approaches

No established unifying theory or paradigm guides AI research. Researchers disagree about many issues.[164] A few of the most long-standing questions that have remained unanswered are these: should artificial intelligence simulate natural intelligence by studying psychology or neurobiology? Or is human biology as irrelevant to AI research as bird biology is to aeronautical engineering?[21] Can intelligent behavior be described using simple, elegant principles (such as logic or optimization)? Or does it necessarily require solving a large number of unrelated problems?[22]

Cybernetics and brain simulation

In the 1940s and 1950s, a number of researchers explored the connection between neurobiology, information theory, and cybernetics. Some of them built machines that used electronic networks to exhibit rudimentary intelligence, such as W. Grey Walter's turtles and the Johns Hopkins Beast. Many of these researchers gathered for meetings of the Teleological Society at Princeton University and the Ratio Club in England.[165] By 1960, this approach was largely abandoned, although elements of it would be revived in the 1980s.

Symbolic

When access to digital computers became possible in the mid-1950s, AI research began to explore the possibility that human intelligence could be reduced to symbol manipulation. The research was centered in three institutions: Carnegie Mellon University, Stanford and MIT, and as described below, each one developed its own style of research. John Haugeland named these symbolic approaches to AI "good old fashioned AI" or "GOFAI".[166] During the 1960s, symbolic approaches had achieved great success at simulating high-level "thinking" in small demonstration programs. Approaches based on cybernetics or artificial neural networks were abandoned or pushed into the background.[167] 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.

Cognitive simulation

Economist Herbert Simon and Allen Newell studied human problem-solving skills and attempted to formalize them, and their work laid the foundations of the field of artificial intelligence, as well as cognitive science, operations research and management science. Their research team used the results of psychological experiments to develop programs that simulated the techniques that people used to solve problems. This tradition, centered at Carnegie Mellon University would eventually culminate in the development of the Soar architecture in the middle 1980s.[168][169]

Logic-based

Unlike Simon and Newell, John McCarthy felt that machines did not need to simulate human thought, but should instead try to find the essence of abstract reasoning and problem-solving, regardless of whether people used the same algorithms.[21] His laboratory at Stanford (SAIL) focused on using formal logic to solve a wide variety of problems, including knowledge representation, planning and learning.[170] Logic was also the focus of the work at the University of Edinburgh and elsewhere in Europe which led to the development of the programming language Prolog and the science of logic programming.[171]

Anti-logic or scruffy

Researchers at MIT (such as Marvin Minsky and Seymour Papert)[172] found that solving difficult problems in vision and natural language processing required ad hoc solutions—they argued that no simple and general principle (like logic) would capture all the aspects of intelligent behavior. Roger Schank described their "anti-logic" approaches as "scruffy" (as opposed to the "neat" paradigms at CMU and Stanford).[22] Commonsense knowledge bases (such as Doug Lenat's Cyc) are an example of "scruffy" AI, since they must be built by hand, one complicated concept at a time.[173]

Knowledge-based

When computers with large memories became available around 1970, researchers from all three traditions began to build knowledge into AI applications.[174] This "knowledge revolution" led to the development and deployment of expert systems (introduced by Edward Feigenbaum), the first truly successful form of AI software.[50] A key component of the system architecture for all expert systems is the knowledge base, which stores facts and rules that illustrate AI.[175] The knowledge revolution was also driven by the realization that enormous amounts of knowledge would be required by many simple AI applications.

Sub-symbolic

By the 1980s, progress in symbolic AI seemed to stall and many believed that symbolic systems would never 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.[23] Sub-symbolic methods manage to approach intelligence without specific representations of knowledge.

Embodied intelligence

This includes embodied, situated, behavior-based, and nouvelle AI. Researchers from the related field of robotics, such as Rodney Brooks, rejected symbolic AI and focused on the basic engineering problems that would allow robots to move and survive.[176] Their work revived the non-symbolic point of view of the early cybernetics researchers of the 1950s and reintroduced the use of control theory in AI. This coincided with the development of the embodied mind thesis in the related field of cognitive science: the idea that aspects of the body (such as movement, perception and visualization) are required for higher intelligence.

Within developmental robotics, developmental learning approaches are elaborated upon to allow robots to accumulate repertoires of novel skills through autonomous self-exploration, social interaction with human teachers, and the use of guidance mechanisms (active learning, maturation, motor synergies, etc.).[177][178][179][180]

Computational intelligence and soft computing

Interest in neural networks and "connectionism" was revived by David Rumelhart and others in the middle of the 1980s.[181] Artificial neural networks are an example of soft computing—they are solutions to problems which cannot be solved with complete logical certainty, and where an approximate solution is often sufficient. Other soft computing approaches to AI include fuzzy systems, Grey system theory, evolutionary computation and many statistical tools. The application of soft computing to AI is studied collectively by the emerging discipline of computational intelligence.[182]

Statistical

Much of traditional GOFAI got bogged down on *ad hoc* patches to symbolic computation that worked on their own toy models but failed to generalize to real-world results. However, around the 1990s, AI researchers adopted sophisticated mathematical tools, such as hidden Markov models (HMM), information theory, and normative Bayesian decision theory to compare or to unify competing architectures. The shared mathematical language permitted a high level of collaboration with more established fields (like mathematics, economics or operations research).[d] Compared with GOFAI, new "statistical learning" techniques such as HMM and neural networks were gaining higher levels of accuracy in many practical domains such as data mining, without necessarily acquiring a semantic understanding of the datasets. The increased successes with real-world data led to increasing emphasis on comparing different approaches against shared test data to see which approach performed best in a broader context than that provided by idiosyncratic toy models; AI research was becoming more scientific. Nowadays results of experiments are often rigorously measurable, and are sometimes (with difficulty) reproducible.[52][183] Different statistical learning techniques have different limitations; for example, basic HMM cannot model the infinite possible combinations of natural language.[184] Critics note that the shift from GOFAI to statistical learning is often also a shift away from explainable AI. In AGI research, some scholars caution against over-reliance on statistical learning, and argue that continuing research into GOFAI will still be necessary to attain general intelligence.[185][186]

Integrating the approaches

Intelligent agent paradigm

An intelligent agent is a system that perceives its environment and takes actions that maximize its chances of success. The simplest intelligent agents are programs that solve specific problems. More complicated agents include human beings and organizations of human beings (such as firms). The paradigm allows researchers to directly compare or even combine different approaches to isolated problems, by asking which agent is best at maximizing a given "goal function". An agent that solves a specific problem can use any approach that works—some agents are symbolic and logical, some are sub-symbolic artificial neural networks and others may use new approaches. The paradigm also gives researchers a common language to communicate with other fields—such as decision theory and economics—that also use concepts of abstract agents. Building a complete agent requires researchers to address realistic problems of integration; for example, because sensory systems give uncertain information about the environment, planning systems must be able to function in the presence of uncertainty. The intelligent agent paradigm became widely accepted during the 1990s.[187]

Agent architectures and cognitive architectures

Researchers have designed systems to build intelligent systems out of interacting intelligent agents in a multi-agent system.[188] A hierarchical control system provides a bridge between sub-symbolic AI at its lowest, reactive levels and traditional symbolic AI at its highest levels, where relaxed time constraints permit planning and world modeling.[189] Some cognitive architectures are custom-built to solve a narrow problem; others, such as Soar, are designed to mimic human cognition and to provide insight into general intelligence. Modern extensions of Soar are hybrid intelligent systems that include both symbolic and sub-symbolic components.[190][191]

Tools

Applications

AI is relevant to any intellectual task.[192] Modern artificial intelligence techniques are pervasive[193] and are too numerous to list here. Frequently, when a technique reaches mainstream use, it is no longer considered artificial intelligence; this phenomenon is described as the AI effect.[194]

High-profile examples of AI include autonomous vehicles (such as drones and self-driving cars), medical diagnosis, creating art (such as poetry), proving mathematical theorems, playing games (such as Chess or Go), search engines (such as Google search), online assistants (such as Siri), image recognition in photographs, spam filtering, predicting flight delays,[195] prediction of judicial decisions,[196] targeting online advertisements, [192][197][198] and energy storage[199]

With social media sites overtaking TV as a source for news for young people and news organizations increasingly reliant on social media platforms for generating distribution,[200] major publishers now use artificial intelligence (AI) technology to post stories more effectively and generate higher volumes of traffic.[201]

AI can also produce Deepfakes, a content-altering technology. ZDNet reports, "It presents something that did not actually occur," Though 88% of Americans believe Deepfakes can cause more harm than good, only 47% of them believe they can be targeted. The boom of election year also opens public discourse to threats of videos of falsified politician media.[202]

Philosophy and ethics

There are three philosophical questions related to AI:

1. Whether artificial general intelligence is possible; whether a machine can solve any problem that a human being can solve using intelligence, or if there are hard limits to what a machine can accomplish.
2. Whether intelligent machines are dangerous; how humans can ensure that machines behave ethically and that they are used ethically.
3. Whether a machine can have a mind, consciousness and mental states in the same sense that human beings do; if a machine can be sentient, and thus deserve certain rights − and if a machine can intentionally cause harm.

The limits of artificial general intelligence

Alan Turing's "polite convention"

One need not decide if a machine can "think"; one need only decide if a machine can act as intelligently as a human being. This approach to the philosophical problems associated with artificial intelligence forms the basis of the Turing test.[203]

The Dartmouth proposal

"Every aspect of learning or any other feature of intelligence can be so precisely described that a machine can be made to simulate it." This conjecture was printed in the proposal for the Dartmouth Conference of 1956.[204]

Newell and Simon's physical symbol system hypothesis

"A physical symbol system has the necessary and sufficient means of general intelligent action." Newell and Simon argue that intelligence consists of formal operations on symbols.[205] Hubert Dreyfus argues that, on the contrary, human expertise depends on unconscious instinct rather than conscious symbol manipulation, and on having a "feel" for the situation, rather than explicit symbolic knowledge. (See Dreyfus' critique of AI.)[207][208]

Gödelian arguments

Gödel himself,[209] John Lucas (in 1961) and Roger Penrose (in a more detailed argument from 1989 onwards) made highly technical arguments that human mathematicians can consistently see the truth of their own "Gödel statements" and therefore have computational abilities beyond that of mechanical Turing machines.[210] However, some people do not agree with the "Gödelian arguments".[211][212][213]

The artificial brain argument

An argument asserting that the brain can be simulated by machines and, because brains exhibit intelligence, these simulated brains must also exhibit intelligence − ergo, machines can be intelligent. Hans Moravec, Ray Kurzweil and others have argued that it is technologically feasible to copy the brain directly into hardware and software, and that such a simulation will be essentially identical to the original.[161]

The AI effect

A hypothesis claiming that machines are *already* intelligent, but observers have failed to recognize it. For example, when Deep Blue beat Garry Kasparov in chess, the machine could be described as exhibiting intelligence. However, onlookers commonly discount the behavior of an artificial intelligence program by arguing that it is not "real" intelligence, with "real" intelligence being in effect defined as whatever behavior machines cannot do.

Ethical machines

Machines with intelligence have the potential to use their intelligence to prevent harm and minimize the risks; they may have the ability to use ethical reasoning to better choose their actions in the world. As such, there is a need for policy making to devise policies for and regulate artificial intelligence and robotics.[214] Research in this area includes machine ethics, artificial moral agents, friendly AI and discussion towards building a human rights framework is also in talks.[215]

Joseph Weizenbaum in *Computer Power and Human Reason* wrote that AI applications cannot, by definition, successfully simulate genuine human empathy and that the use of AI technology in fields such as customer service or psychotherapy[217] was deeply misguided. Weizenbaum was also bothered that AI researchers (and some philosophers) were willing to view the human mind as nothing more than a computer program (a position now known as computationalism). To Weizenbaum these points suggest that AI research devalues human life.[218]

Artificial moral agents

Wendell Wallach introduced the concept of artificial moral agents (AMA) in his book *Moral Machines*[219] For Wallach, AMAs have become a part of the research landscape of artificial intelligence as guided by its two central questions which he identifies as "Does Humanity Want Computers Making Moral Decisions"[220] and "Can (Ro)bots Really Be Moral".[221] For Wallach, the question is not centered on the issue of *whether* machines can demonstrate the equivalent of moral behavior, unlike the *constraints* which society may place on the development of AMAs.[222]

Machine ethics

The field of machine ethics is concerned with giving machines ethical principles, or a procedure for discovering a way to resolve the ethical dilemmas they might encounter, enabling them to function in an ethically responsible manner through their own ethical decision making.[223] The field was delineated in the AAAI Fall 2005 Symposium on Machine Ethics: "Past research concerning the relationship between technology and ethics has largely focused on responsible and irresponsible use of technology by human beings, with a few people being interested in how human beings ought to treat machines. In all cases, only human beings have engaged in ethical reasoning. The time has come for adding an ethical dimension to at least some machines. Recognition of the ethical ramifications of behavior involving machines, as well as recent and potential developments in machine autonomy, necessitate this. In contrast to computer hacking, software property issues, privacy issues and other topics normally ascribed to computer ethics, machine ethics is concerned with the behavior of machines towards human users and other machines. Research in machine ethics is key to alleviating concerns with autonomous systems—it could be argued that the notion of autonomous machines without such a dimension is at the root of all fear concerning machine intelligence. Further, investigation of machine ethics could enable the discovery of problems with current ethical theories, advancing our thinking about Ethics."[224] Machine ethics is sometimes referred to as machine morality, computational ethics or computational morality. A variety of perspectives of this nascent field can be found in the collected edition "Machine Ethics"[223] that stems from the AAAI Fall 2005 Symposium on Machine Ethics.[224]

Malevolent and friendly AI

Political scientist Charles T. Rubin believes that AI can be neither designed nor guaranteed to be benevolent.[225] He 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 be sympathetic to our system of morality, which has evolved along with our particular biology (which AIs would not share). Hyper-intelligent software may not necessarily decide to support the continued existence of humanity and would be extremely difficult to stop. This topic has also recently begun to be discussed in academic publications as a real source of risks to civilization, humans, and planet Earth.

One proposal to deal with this is to ensure that the first generally intelligent AI is 'Friendly AI' and will be able to control subsequently developed AIs. Some question whether this kind of check could actually remain in place.

Leading AI researcher Rodney Brooks writes, "I think it is a mistake to be worrying about us developing malevolent AI anytime in the next few hundred years. I think the worry stems from a fundamental error in not distinguishing the difference between the very real recent advances in a particular aspect of AI and the enormity and complexity of building sentient volitional intelligence."[226]

Lethal autonomous weapons are of concern. Currently, 50+ countries are researching battlefield robots, including the United States, China, Russia, and the United Kingdom. Many people concerned about risk from superintelligent AI also want to limit the use of artificial soldiers and drones.[227]

Machine consciousness, sentience and mind

If an AI system replicates all key aspects of human intelligence, will that system also be sentient—will it have a mind which has conscious experiences? This question is closely related to the philosophical problem as to the nature of human consciousness, generally referred to as the hard problem of consciousness.

Consciousness

David Chalmers identified two problems in understanding the mind, which he named the "hard" and "easy" problems of consciousness.[228] 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, consider what happens when a person is shown a color swatch and identifies it, saying "it's red". The easy problem only requires understanding the machinery in the brain that makes it possible for a person to know that the color swatch is red. The hard problem is that people also know something else—they also know *what red looks like*. (Consider that a person born blind can know that something is red without knowing what red looks like.)[e] Everyone knows subjective experience exists, because they do it every day (e.g., all sighted people know what red looks like). The hard problem is explaining how the brain creates it, why it exists, and how it is different from knowledge and other aspects of the brain.

Computationalism and functionalism

Computationalism is the position in the philosophy of mind that the human mind or the human brain (or both) is an information processing system and that thinking is a form of computing.[229] 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.

Strong AI hypothesis

The philosophical position that John Searle has named "strong AI" states: "The appropriately programmed computer with the right inputs and outputs would thereby have a mind in exactly the same sense human beings have minds."[231] Searle counters this assertion with his Chinese room argument, which asks us to look *inside* the computer and try to find where the "mind" might be.[232]

Robot rights

If a machine can be created that has intelligence, could it also *feel*? If it can feel, does it have the same rights as a human? This issue, now known as "robot rights", is currently being considered by, for example, California's Institute for the Future, although many critics believe that the discussion is premature.[233] Some critics of transhumanism argue that any hypothetical robot rights would lie on a spectrum with animal rights and human rights. [234] The subject is profoundly discussed in the 2010 documentary film *Plug & Pray*,[235] and many sci fi media such as Star Trek Next Generation, with the character of Commander Data, who fought being disassembled for research, and wanted to "become human", and the robotic holograms in Voyager.

Superintelligence

Are there limits to how intelligent machines—or human-machine hybrids—can be? 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.[155]

Technological singularity

If research into Strong AI 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.[236] The new intelligence could thus increase exponentially and dramatically surpass humans. Science fiction writer Vernor Vinge named this scenario "singularity".[237] Technological singularity is when accelerating progress in technologies will cause a runaway effect wherein artificial intelligence will exceed human intellectual capacity and control, thus radically changing or even ending civilization. Because the capabilities of such an intelligence may be impossible to comprehend, the technological singularity is an occurrence beyond which events are unpredictable or even unfathomable.[237][155]

Ray Kurzweil has used Moore's law (which describes the relentless exponential improvement in digital technology) to calculate that desktop computers will have the same processing power as human brains by the year 2029 and predicts that the singularity will occur in 2045.[237]

Transhumanism

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.[238] This idea, called transhumanism, has roots in Aldous Huxley and Robert Ettinger.

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.[239]

Impact

The long-term economic effects of AI are uncertain. 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.[240] A February 2020 European Union white paper on artificial intelligence advocated for artificial intelligence for economic benefits, including "improving healthcare (e.g. making diagnosis more precise, enabling better prevention of diseases), increasing the efficiency of farming, contributing to climate change mitigation and adaptation, [and] improving the efficiency of production systems through predictive maintenance", while acknowledging potential risks.[193]

The relationship between automation and employment is complicated. While automation eliminates old jobs, it also creates new jobs through micro-economic and macro-economic effects.[241] 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".[242] 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".[243][244][245] 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.[246] Author Martin Ford and others go further and argue that many jobs are routine, repetitive and (to an AI) predictable; Ford warns that these jobs may be automated in the next couple of decades, and that many of the new jobs may not be "accessible to people with average capability", even with retraining. Economists point out that in the past technology has tended to increase rather than reduce total employment, but acknowledge that "we're in uncharted territory" with AI.[33]

The potential negative effects of AI and automation were a major issue for Andrew Yang's 2020 presidential campaign in the United States.[247] Irakli Beridze, Head of the Centre for Artificial Intelligence and Robotics at UNICRI, United Nations, has expressed that "I think the dangerous applications for AI, from my point of view, would be criminals or large terrorist organizations using it to disrupt large processes or simply do pure harm. [Terrorists could cause harm] via digital warfare, or it could be a combination of robotics, drones, with AI and other things as well that could be really dangerous. And, of course, other risks come from things like job losses. If we have massive numbers of people losing jobs and don't find a solution, it will be extremely dangerous. Things like lethal autonomous weapons systems should be properly governed — otherwise there's massive potential of misuse."[248]

Risks of narrow AI

Widespread use of artificial intelligence could have unintended consequences that are dangerous or undesirable. Scientists from the Future of Life Institute, among others, described some short-term research goals to see how AI influences the economy, the laws and ethics that are involved with AI and how to minimize AI security risks. In the long-term, the scientists have proposed to continue optimizing function while minimizing possible security risks that come along with new technologies.[249]

Some are concerned about algorithmic bias, that AI programs may unintentionally become biased after processing data that exhibits bias.[250] Algorithms already have numerous applications in legal systems. 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 average COMPAS-assigned recidivism risk level of black defendants is significantly higher than the average COMPAS-assigned risk level of white defendants.[251]

Risks of general AI

Physicist Stephen Hawking, Microsoft founder Bill Gates, and SpaceX founder Elon Musk have expressed concerns about the possibility that AI could evolve to the point that humans could not control it, with Hawking theorizing that this could "spell the end of the human race".[252][253][254]

> The development of full artificial intelligence could spell the end of the human race. Once humans develop artificial intelligence, it will take off on its own and redesign itself at an ever-increasing rate. Humans, who are limited by slow biological evolution, couldn't compete and would be superseded.
>
> — Stephen Hawking[255]

In his book *Superintelligence*, philosopher Nick Bostrom provides an argument that artificial intelligence will pose a threat to humankind. He 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—one example is an AI told to compute as many digits of pi as possible—it might harm humanity in order to acquire more resources or prevent itself from being shut down, ultimately to better achieve its goal. Bostrom also emphasizes the difficulty of fully conveying humanity's values to an advanced AI. He uses the hypothetical example of giving an AI the goal to make humans smile to illustrate a misguided attempt. If the AI in that scenario were to become superintelligent, Bostrom argues, it may resort to methods that most humans would find horrifying, such as inserting "electrodes into the facial muscles of humans to cause constant, beaming grins" because that would be an efficient way to achieve its goal of making humans smile.[256] In his book *Human Compatible*, AI researcher Stuart J. Russell echoes some of Bostrom's concerns while also proposing an approach to developing provably beneficial machines focused on uncertainty and deference to humans,[257]:173 possibly involving inverse reinforcement learning.[257]:191–193

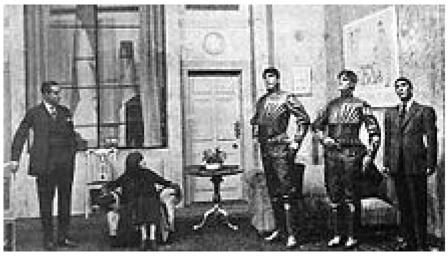
Concern over risk from artificial intelligence has led to some high-profile donations and investments. A group of prominent tech titans including Peter Thiel, Amazon Web Services and Musk have committed $1 billion to OpenAI, a nonprofit company aimed at championing responsible AI development.[258] The opinion of experts within the field of artificial intelligence is mixed, with sizable fractions both concerned and unconcerned by risk from eventual superhumanly-capable AI.[259] Other technology industry leaders believe that artificial intelligence is helpful in its current form and will continue to assist humans. Oracle CEO Mark Hurd has stated that AI "will actually create more jobs, not less jobs" as humans will be needed to manage AI systems.[260] Facebook CEO Mark Zuckerberg believes AI will "unlock a huge amount of positive things," such as curing disease and increasing the safety of autonomous cars.[261] In January 2015, Musk donated $10 million to the Future of Life Institute to fund research on understanding AI decision making. The goal of the institute is to "grow wisdom with which we manage" the growing power of technology. Musk also funds companies developing artificial intelligence such as DeepMind and Vicarious to "just keep an eye on what's going on with artificial intelligence.[262] I think there is potentially a dangerous outcome there."[263][264]

For the danger of uncontrolled advanced AI to be realized, the hypothetical AI would have to overpower or out-think all of humanity, which a minority of experts argue is a possibility far enough in the future to not be worth researching.[265][266] Other counterarguments revolve around humans being either intrinsically or convergently valuable from the perspective of an artificial intelligence.[267]

Regulation

The regulation of artificial intelligence is the development of public sector policies and laws for promoting and regulating artificial intelligence (AI);[268][269] it is therefore related to the broader regulation of algorithms. The regulatory and policy landscape for AI is an emerging issue in jurisdictions globally, including in the European Union.[270] Regulation is considered necessary to both encourage AI and manage associated risks.[271][272] Regulation of AI through mechanisms such as review boards can also be seen as social means to approach the AI control problem.[273]

In fiction



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

Thought-capable artificial beings appeared as storytelling devices since antiquity,[35] and have been a persistent theme in science fiction.

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.[274]

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;[275] 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.[276]

Transhumanism (the merging of humans and machines) is explored in the manga *Ghost in the Shell* and the science-fiction series *Dune*. In the 1980s, artist Hajime Sorayama's Sexy Robots series were painted and published in Japan depicting the actual organic human form with lifelike muscular metallic skins and later "the Gynoids" book followed that was used by or influenced movie makers including George Lucas and other creatives. Sorayama never considered these organic robots to be real part of nature but always an unnatural product of the human mind, a fantasy existing in the mind even when realized in actual form.

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.[277]

See also

- *A.I. Rising*
- AI control problem
- Artificial intelligence arms race
- Behavior selection algorithm

- Business process automation
- Case-based reasoning
- Citizen Science
- Emergent algorithm
- Female gendering of AI technologies
- Glossary of artificial intelligence
- Regulation of artificial intelligence
- Robotic process automation
- Universal basic income
- Weak AI

Explanatory notes

a. The act of doling out rewards can itself be formalized or automated into a "reward function".
b. Terminology varies; see algorithm characterizations.
c. Adversarial vulnerabilities can also result in nonlinear systems, or from non-pattern perturbations. Some systems are so brittle that changing a single adversarial pixel predictably induces misclassification.
d. While such a "victory of the neats" may be a consequence of the field becoming more mature, AIMA states that in practice both neat and scruffy approaches continue to be necessary in AI research.
e. This is based on Mary's Room, a thought experiment first proposed by Frank Jackson in 1982

References

1. Poole, Mackworth & Goebel 1998, p. 1 (http://people.cs.ubc.ca/~poole/ci/ch1.pdf).
2. Russell & Norvig 2003, p. 55.
3. Definition of AI as the study of intelligent agents:
   - Poole, Mackworth & Goebel (1998), which provides the version that is used in this article. These authors use the term "computational intelligence" as a synonym for artificial intelligence.[1]
   - Russell & Norvig (2003) (who prefer the term "rational agent") and write "The whole-agent view is now widely accepted in the field".[2]
   - Nilsson 1998
   - Legg & Hutter 2007
4. Russell & Norvig 2009, p. 2.
5. McCorduck 2004, p. 204
6. Maloof, Mark. "Artificial Intelligence: An Introduction, p. 37" (http://people.cs.georgetown.edu/~maloof/cosc270.f17/cosc270-intro-handout.pdf) (PDF). *georgetown.edu*. Archived (https://web.archive.org/web/20180825070509/http://people.cs.georgetown.edu/~maloof/cosc270.f17/cosc270-intro-handout.pdf) (PDF) from the original on 25 August 2018.
7. "How AI Is Getting Groundbreaking Changes In Talent Management And HR Tech" (https://hackernoon.com/how-ai-is-getting-groundbreaking-changes-in-talent-management-and-hr-tech-d24ty3zzd). Hackernoon. Archived (https://web.archive.org/web/20190911083500/https://hackernoon.com/how-ai-is-getting-groundbreaking-changes-in-talent-management-and-hr-tech-d24ty3zzd) from the original on 11 September 2019. Retrieved 14 February 2020.
8. Schank, Roger C. (1991). "Where's the AI". *AI magazine*. Vol. 12 no. 4. p. 38.
9. Russell & Norvig 2009.
10. "AlphaGo – Google DeepMind" (https://deepmind.com/alpha-go.html). Archived (https://web.archive.org/web/20160310191926/https://www.deepmind.com/alpha-go.html) from the original on 10 March 2016.

11. Allen, Gregory (April 2020). "Department of Defense Joint AI Center - Understanding AI Technology" (https://www.ai.mil/docs/Understanding%20AI%20Technology.pdf) (PDF). *AI.mil - The official site of the Department of Defense Joint Artificial Intelligence Center*. Archived (https://web.archive.org/web/20200421074921/https://www.ai.mil/docs/Understanding%20AI%20Technology.pdf) (PDF) from the original on 21 April 2020. Retrieved 25 April 2020.
12. Optimism of early AI: * Herbert Simon quote: Simon 1965, p. 96 quoted in Crevier 1993, p. 109. * Marvin Minsky quote: Minsky 1967, p. 2 quoted in Crevier 1993, p. 109.
13. Boom of the 1980s: rise of expert systems, Fifth Generation Project, Alvey, MCC, SCI: * 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
14. First AI Winter, Mansfield Amendment, Lighthill report * Crevier 1993, pp. 115–117 * Russell & Norvig 2003, p. 22 * NRC 1999, pp. 212–213 * Howe 1994 * Newquist 1994, pp. 189–201
15. Second AI winter: * McCorduck 2004, pp. 430–435 * Crevier 1993, pp. 209–210 * NRC 1999, pp. 214–216 * Newquist 1994, pp. 301–318
16. AI becomes hugely successful in the early 21st century * Clark 2015
17. Pamela McCorduck (2004, p. 424) writes of "the rough shattering of AI in subfields—vision, natural language, decision theory, genetic algorithms, robotics ... and these with own sub-subfield—that would hardly have anything to say to each other."
18. 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 * Nilsson 1998
19. Kolata 1982.
20. Maker 2006.
21. Biological intelligence vs. intelligence in general:
    - Russell & Norvig 2003, pp. 2–3, who make the analogy with aeronautical engineering.
    - McCorduck 2004, pp. 100–101, who writes that there are "two major branches of artificial intelligence: one aimed at producing intelligent behavior regardless of how it was accomplished, and the other aimed at modeling intelligent processes found in nature, particularly human ones."
    - Kolata 1982, a paper in *Science*, which describes McCarthy's indifference to biological models. Kolata quotes McCarthy as writing: "This is AI, so we don't care if it's psychologically real".[19] McCarthy recently reiterated his position at the AI@50 conference where he said "Artificial intelligence is not, by definition, simulation of human intelligence".[20].
22. Neats vs. scruffies: * McCorduck 2004, pp. 421–424, 486–489 * Crevier 1993, p. 168 * Nilsson 1983, pp. 10–11
23. Symbolic vs. sub-symbolic AI: * Nilsson (1998, p. 7), who uses the term "sub-symbolic".
24. General intelligence (strong AI) is discussed in popular introductions to AI: * Kurzweil 1999 and Kurzweil 2005
25. See the Dartmouth proposal, under Philosophy, below.
26. McCorduck 2004, p. 34.
27. McCorduck 2004, p. xviii.
28. McCorduck 2004, p. 3.
29. McCorduck 2004, pp. 340–400.

30. This is a central idea of Pamela McCorduck's *Machines Who Think*. She writes:

    - "I like to think of artificial intelligence as the scientific apotheosis of a venerable cultural tradition."[26]
    - "Artificial intelligence in one form or another is an idea that has pervaded Western intellectual history, a dream in urgent need of being realized."[27]
    - "Our history is full of attempts—nutty, eerie, comical, earnest, legendary and real—to make artificial intelligences, to reproduce what is the essential us—bypassing the ordinary means. Back and forth between myth and reality, our imaginations supplying what our workshops couldn't, we have engaged for a long time in this odd form of self-reproduction."[28]

    She traces the desire back to its Hellenistic roots and calls it the urge to "forge the Gods."[29]
31. "Stephen Hawking believes AI could be mankind's last accomplishment" (https://betanews.com/2016/10/21/artificial-intelligence-stephen-hawking/). *BetaNews*. 21 October 2016. Archived (https://web.archive.org/web/20170828183930/https://betanews.com/2016/10/21/artificial-intelligence-stephen-hawking/) from the original on 28 August 2017.
32. Lombardo P, Boehm I, Nairz K (2020). "RadioComics – Santa Claus and the future of radiology" (https://doi.org/10.1016%2Fj.ejrad.2019.108771). *Eur J Radiol*. **122** (1): 108771. doi:10.1016/j.ejrad.2019.108771 (https://doi.org/10.1016%2Fj.ejrad.2019.108771). PMID 31835078 (https://pubmed.ncbi.nlm.nih.gov/31835078).
33. Ford, Martin; Colvin, Geoff (6 September 2015). "Will robots create more jobs than they destroy?" (https://www.theguardian.com/technology/2015/sep/06/will-robots-create-destroy-jobs). *The Guardian*. Archived (https://web.archive.org/web/20180616204119/https://www.theguardian.com/technology/2015/sep/06/will-robots-create-destroy-jobs) from the original on 16 June 2018. Retrieved 13 January 2018.
34. AI applications widely used behind the scenes: * Russell & Norvig 2003, p. 28 * Kurzweil 2005, p. 265 * NRC 1999, pp. 216–222 * Newquist 1994, pp. 189–201
35. AI in myth: * McCorduck 2004, pp. 4–5 * Russell & Norvig 2003, p. 939
36. AI in early science fiction. * McCorduck 2004, pp. 17–25
37. Formal reasoning: * Berlinski, David (2000). *The Advent of the Algorithm* (https://archive.org/details/adventofalgorith0000berl). Harcourt Books. ISBN 978-0-15-601391-8. OCLC 46890682 (https://www.worldcat.org/oclc/46890682). Archived (https://web.archive.org/web/20200726215744/https://archive.org/details/adventofalgorith0000berl) from the original on 26 July 2020. Retrieved 22 August 2020.
38. 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
39. Russell & Norvig 2009, p. 16.
40. Dartmouth conference: * McCorduck 2004, pp. 111–136 * Crevier 1993, pp. 47–49, who writes "the conference is generally recognized as the official birthdate of the new science." * Russell & Norvig 2003, p. 17, who call the conference "the birth of artificial intelligence." * NRC 1999, pp. 200–201
41. McCarthy, John (1988). "Review of *The Question of Artificial Intelligence*". *Annals of the History of Computing*. **10** (3): 224–229., collected in McCarthy, John (1996). "10. Review of *The Question of Artificial Intelligence*". *Defending AI Research: A Collection of Essays and Reviews*. CSLI., p. 73, "[O]ne of the reasons for inventing the term "artificial intelligence" was to escape association with "cybernetics". Its concentration on analog feedback seemed misguided, and I wished to avoid having either to accept Norbert (not Robert) Wiener as a guru or having to argue with him."
42. Hegemony of the Dartmouth conference attendees: * Russell & Norvig 2003, p. 17, who write "for the next 20 years the field would be dominated by these people and their students." * McCorduck 2004, pp. 129–130

43. Russell & Norvig 2003, p. 18.
44. Schaeffer J. (2009) Didn't Samuel Solve That Game?. In: One Jump Ahead. Springer, Boston, MA
45. Samuel, A. L. (July 1959). "Some Studies in Machine Learning Using the Game of Checkers". *IBM Journal of Research and Development*. **3** (3): 210–229. CiteSeerX 10.1.1.368.2254 (https://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.368.2254). doi:10.1147/rd.33.0210 (https://doi.org/10.1147%2Frd.33.0210).
46. "Golden years" of AI (successful symbolic reasoning programs 1956–1973): * McCorduck 2004, pp. 243–252 * Crevier 1993, pp. 52–107 * Moravec 1988, p. 9 * Russell & Norvig 2003, pp. 18–21 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.
47. DARPA pours money into undirected pure research into AI during the 1960s: * McCorduck 2004, p. 131 * Crevier 1993, pp. 51, 64–65 * NRC 1999, pp. 204–205
48. AI in England: * Howe 1994
49. Lighthill 1973.
50. Expert systems: * ACM 1998, I.2.1 * 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
51. Mead, Carver A.; Ismail, Mohammed (8 May 1989). *Analog VLSI Implementation of Neural Systems* (https://web.archive.org/web/20191106154442/http://fennetic.net/irc/Christopher%20R.%20Carroll%20Carver%20Mead%20Mohammed%20Ismail%20Analog%20VLSI%20Implementation%20of%20Neural%20Systems.pdf) (PDF). The Kluwer International Series in Engineering and Computer Science. **80**. Norwell, MA: Kluwer Academic Publishers. doi:10.1007/978-1-4613-1639-8 (https://doi.org/10.1007%2F978-1-4613-1639-8). ISBN 978-1-4613-1639-8. Archived from the original (http://fennetic.net/irc/Christopher%20R.%20Carroll%20Carver%20Mead%20Mohammed%20Ismail%20Analog%20VLSI%20Implementation%20of%20Neural%20Systems.pdf) (PDF) on 6 November 2019. Retrieved 24 January 2020.
52. Formal methods are now preferred ("Victory of the neats"): * Russell & Norvig 2003, pp. 25–26 * McCorduck 2004, pp. 486–487
53. McCorduck 2004, pp. 480–483.
54. Markoff 2011.
55. "Ask the AI experts: What's driving today's progress in AI?" (https://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/ask-the-ai-experts-whats-driving-todays-progress-in-ai). *McKinsey & Company*. Archived (https://web.archive.org/web/20180413190018/https://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/ask-the-ai-experts-whats-driving-todays-progress-in-ai) from the original on 13 April 2018. Retrieved 13 April 2018.
56. Administrator. "Kinect's AI breakthrough explained" (http://www.i-programmer.info/news/105-artificial-intelligence/2176-kinects-ai-breakthrough-explained.html). *i-programmer.info*. Archived (https://web.archive.org/web/20160201031242/http://www.i-programmer.info/news/105-artificial-intelligence/2176-kinects-ai-breakthrough-explained.html) from the original on 1 February 2016.
57. Rowinski, Dan (15 January 2013). "Virtual Personal Assistants & The Future Of Your Smartphone [Infographic]" (http://readwrite.com/2013/01/15/virtual-personal-assistants-the-future-of-your-smartphone-infographic). *ReadWrite*. Archived (https://web.archive.org/web/20151222083034/http://readwrite.com/2013/01/15/virtual-personal-assistants-the-future-of-your-smartphone-infographic) from the original on 22 December 2015.
58. "Artificial intelligence: Google's AlphaGo beats Go master Lee Se-dol" (https://www.bbc.com/news/technology-35785875). *BBC News*. 12 March 2016. Archived (https://web.archive.org/web/20160826103910/http://www.bbc.com/news/technology-35785875) from the original on 26 August 2016. Retrieved 1 October 2016.

59. Metz, Cade (27 May 2017). "After Win in China, AlphaGo's Designers Explore New AI" (https://www.wired.com/2017/05/win-china-alphagos-designers-explore-new-ai/). *Wired*. Archived (https://web.archive.org/web/20170602234726/https://www.wired.com/2017/05/win-china-alphagos-designers-explore-new-ai/) from the original on 2 June 2017.
60. "World's Go Player Ratings" (http://www.goratings.org/). May 2017. Archived (https://web.archive.org/web/20170401123616/https://www.goratings.org/) from the original on 1 April 2017.
61. "柯洁迎19岁生日 雄踞人类世界排名第一已两年" (http://sports.sina.com.cn/go/2016-08-02/doc-ifxunyya3020238.shtml) (in Chinese). May 2017. Archived (https://web.archive.org/web/20170811222849/http://sports.sina.com.cn/go/2016-08-02/doc-ifxunyya3020238.shtml) from the original on 11 August 2017.
62. Clark, Jack (8 December 2015). "Why 2015 Was a Breakthrough Year in Artificial Intelligence" (https://www.bloomberg.com/news/articles/2015-12-08/why-2015-was-a-breakthrough-year-in-artificial-intelligence). *Bloomberg News*. Archived (https://web.archive.org/web/20161123053855/https://www.bloomberg.com/news/articles/2015-12-08/why-2015-was-a-breakthrough-year-in-artificial-intelligence) from the original on 23 November 2016. Retrieved 23 November 2016. "After a half-decade of quiet breakthroughs in artificial intelligence, 2015 has been a landmark year. Computers are smarter and learning faster than ever."
63. "Reshaping Business With Artificial Intelligence" (https://sloanreview.mit.edu/projects/reshaping-business-with-artificial-intelligence/). *MIT Sloan Management Review*. Archived (https://web.archive.org/web/20180519171905/https://sloanreview.mit.edu/projects/reshaping-business-with-artificial-intelligence/) from the original on 19 May 2018. Retrieved 2 May 2018.
64. Lorica, Ben (18 December 2017). "The state of AI adoption" (https://www.oreilly.com/ideas/the-state-of-ai-adoption). *O'Reilly Media*. Archived (https://web.archive.org/web/20180502140700/https://www.oreilly.com/ideas/the-state-of-ai-adoption) from the original on 2 May 2018. Retrieved 2 May 2018.
65. Allen, Gregory (6 February 2019). "Understanding China's AI Strategy" (https://www.cnas.org/publications/reports/understanding-chinas-ai-strategy). *Center for a New American Security*. Archived (https://web.archive.org/web/20190317004017/https://www.cnas.org/publications/reports/understanding-chinas-ai-strategy) from the original on 17 March 2019.
66. "Review | How two AI superpowers – the U.S. and China – battle for supremacy in the field" (https://www.washingtonpost.com/outlook/in-the-race-for-supremacy-in-artificial-intelligence-its-us-innovation-vs-chinese-ambition/2018/11/02/013e0030-b08c-11e8-aed9-001309990777_story.html). *Washington Post*. 2 November 2018. Archived (https://web.archive.org/web/20181104044639/https://www.washingtonpost.com/outlook/in-the-race-for-supremacy-in-artificial-intelligence-its-us-innovation-vs-chinese-ambition/2018/11/02/013e0030-b08c-11e8-aed9-001309990777_story.html) from the original on 4 November 2018. Retrieved 4 November 2018.
67. at 10:11, Alistair Dabbs 22 Feb 2019. "Artificial Intelligence: You know it isn't real, yeah?" (https://www.theregister.co.uk/2019/02/22/artificial_intelligence_you_know_it_isnt_real_yeah/). *www.theregister.co.uk*. Archived (https://web.archive.org/web/20200521215046/https://www.theregister.co.uk/2019/02/22/artificial_intelligence_you_know_it_isnt_real_yeah/) from the original on 21 May 2020. Retrieved 22 August 2020.
68. "Stop Calling it Artificial Intelligence" (https://joshworth.com/stop-calling-in-artificial-intelligence/). Archived (https://web.archive.org/web/20191202011223/http://joshworth.com/stop-calling-in-artificial-intelligence/) from the original on 2 December 2019. Retrieved 1 December 2019.
69. "AI isn't taking over the world – it doesn't exist yet" (https://www.gbgplc.com/inside/ai/). *GBG Global website*. Archived (https://web.archive.org/web/20200811091659/https://www.gbgplc.com/inside/ai/) from the original on 11 August 2020. Retrieved 22 August 2020.
70. Kaplan, Andreas; Haenlein, Michael (1 January 2019). "Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence". *Business Horizons*. **62** (1): 15–25. doi:10.1016/j.bushor.2018.08.004 (https://doi.org/10.1016%2Fj.bushor.2018.08.004).
71. Domingos 2015, Chapter 5.

72. Domingos 2015, Chapter 7.
73. Lindenbaum, M., Markovitch, S., & Rusakov, D. (2004). Selective sampling for nearest neighbor classifiers. Machine learning, 54(2), 125–152.
74. Domingos 2015, Chapter 1.
75. Intractability and efficiency and the combinatorial explosion: * Russell & Norvig 2003, pp. 9, 21–22
76. Domingos 2015, Chapter 2, Chapter 3.
77. Hart, P. E.; Nilsson, N. J.; Raphael, B. (1972). "Correction to "A Formal Basis for the Heuristic Determination of Minimum Cost Paths" ". *SIGART Newsletter* (37): 28–29. doi:10.1145/1056777.1056779 (https://doi.org/10.1145%2F1056777.1056779). S2CID 6386648 (https://api.semanticscholar.org/CorpusID:6386648).
78. Domingos 2015, Chapter 2, Chapter 4, Chapter 6.
79. "Can neural network computers learn from experience, and if so, could they ever become what we would call 'smart'?" (https://www.scientificamerican.com/article/can-neural-network-comput/). *Scientific American*. 2018. Archived (https://web.archive.org/web/20180325045717/https://www.scientificamerican.com/article/can-neural-network-comput/) from the original on 25 March 2018. Retrieved 24 March 2018.
80. Domingos 2015, Chapter 6, Chapter 7.
81. Domingos 2015, p. 286.
82. "Single pixel change fools AI programs" (https://www.bbc.com/news/technology-41845878). *BBC News*. 3 November 2017. Archived (https://web.archive.org/web/20180322011306/http://www.bbc.com/news/technology-41845878) from the original on 22 March 2018. Retrieved 12 March 2018.
83. "AI Has a Hallucination Problem That's Proving Tough to Fix" (https://www.wired.com/story/ai-has-a-hallucination-problem-thats-proving-tough-to-fix/). *WIRED*. 2018. Archived (https://web.archive.org/web/20180312024533/https://www.wired.com/story/ai-has-a-hallucination-problem-thats-proving-tough-to-fix/) from the original on 12 March 2018. Retrieved 12 March 2018.
84. Matti, D.; Ekenel, H. K.; Thiran, J. P. (2017). *Combining LiDAR space clustering and convolutional neural networks for pedestrian detection*. *2017 14th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*. pp. 1–6. arXiv:1710.06160 (https://arxiv.org/abs/1710.06160). doi:10.1109/AVSS.2017.8078512 (https://doi.org/10.1109%2FAVSS.2017.8078512). ISBN 978-1-5386-2939-0. S2CID 2401976 (https://api.semanticscholar.org/CorpusID:2401976).
85. Ferguson, Sarah; Luders, Brandon; Grande, Robert C.; How, Jonathan P. (2015). *Real-Time Predictive Modeling and Robust Avoidance of Pedestrians with Uncertain, Changing Intentions*. *Algorithmic Foundations of Robotics XI*. Springer Tracts in Advanced Robotics. **107**. Springer, Cham. pp. 161–177. arXiv:1405.5581 (https://arxiv.org/abs/1405.5581). doi:10.1007/978-3-319-16595-0_10 (https://doi.org/10.1007%2F978-3-319-16595-0_10). ISBN 978-3-319-16594-3. S2CID 8681101 (https://api.semanticscholar.org/CorpusID:8681101).
86. "Cultivating Common Sense | DiscoverMagazine.com" (https://web.archive.org/web/20180325045222/http://discovermagazine.com/2017/april-2017/cultivating-common-sense). *Discover Magazine*. 2017. Archived from the original (http://discovermagazine.com/2017/april-2017/cultivating-common-sense) on 25 March 2018. Retrieved 24 March 2018.
87. Davis, Ernest; Marcus, Gary (24 August 2015). "Commonsense reasoning and commonsense knowledge in artificial intelligence" (https://cacm.acm.org/magazines/2015/9/191169-commonsense-reasoning-and-commonsense-knowledge-in-artificial-intelligence/). *Communications of the ACM*. **58** (9): 92–103. doi:10.1145/2701413 (https://doi.org/10.1145%2F2701413). S2CID 13583137 (https://api.semanticscholar.org/CorpusID:13583137). Archived (https://web.archive.org/web/20200822201550/https://cacm.acm.org/magazines/2015/9/191169-commonsense-reasoning-and-commonsense-knowledge-in-artificial-intelligence/fulltext) from the original on 22 August 2020. Retrieved 6 April 2020.

88. Winograd, Terry (January 1972). "Understanding natural language". *Cognitive Psychology*. **3** (1): 1–191. doi:10.1016/0010-0285(72)90002-3 (https://doi.org/10.1016%2F0010-0285%2872%2990002-3).
89. "Don't worry: Autonomous cars aren't coming tomorrow (or next year)" (http://autoweek.com/article/technology/fully-autonomous-vehicles-are-more-decade-down-road). *Autoweek*. 2016. Archived (https://web.archive.org/web/20180325052230/http://autoweek.com/article/technology/fully-autonomous-vehicles-are-more-decade-down-road) from the original on 25 March 2018. Retrieved 24 March 2018.
90. Knight, Will (2017). "Boston may be famous for bad drivers, but it's the testing ground for a smarter self-driving car" (https://www.technologyreview.com/s/608871/finally-a-driverless-car-with-some-common-sense/). *MIT Technology Review*. Archived (https://web.archive.org/web/20200822201548/https://www.technologyreview.com/2017/09/20/149046/finally-a-driverless-car-with-some-common-sense/) from the original on 22 August 2020. Retrieved 27 March 2018.
91. Prakken, Henry (31 August 2017). "On the problem of making autonomous vehicles conform to traffic law" (https://doi.org/10.1007%2Fs10506-017-9210-0). *Artificial Intelligence and Law*. **25** (3): 341–363. doi:10.1007/s10506-017-9210-0 (https://doi.org/10.1007%2Fs10506-017-9210-0).
92. Lieto, Antonio (May 2018). "The knowledge level in cognitive architectures: Current limitations and possible developments". *Cognitive Systems Research*. **48**: 39–55. doi:10.1016/j.cogsys.2017.05.001 (https://doi.org/10.1016%2Fj.cogsys.2017.05.001). hdl:2318/1665207 (https://hdl.handle.net/2318%2F1665207). S2CID 206868967 (https://api.semanticscholar.org/CorpusID:206868967).
93. 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
94. 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
95. Psychological evidence of sub-symbolic reasoning: * Wason & Shapiro (1966) showed that people do poorly on completely abstract problems, but if the problem is restated to allow the use of intuitive social intelligence, performance dramatically improves. (See Wason selection task) * Kahneman, Slovic & Tversky (1982) have shown that people are terrible at elementary problems that involve uncertain reasoning. (See list of cognitive biases for several examples). * Lakoff & Núñez (2000) have controversially argued that even our skills at mathematics depend on knowledge and skills that come from "the body", i.e. sensorimotor and perceptual skills. (See Where Mathematics Comes From)
96. Knowledge representation: * ACM 1998, I.2.4, * Russell & Norvig 2003, pp. 320–363, * Poole, Mackworth & Goebel 1998, pp. 23–46, 69–81, 169–196, 235–277, 281–298, 319–345, * Luger & Stubblefield 2004, pp. 227–243, * Nilsson 1998, chpt. 18
97. Knowledge engineering: * Russell & Norvig 2003, pp. 260–266, * Poole, Mackworth & Goebel 1998, pp. 199–233, * Nilsson 1998, chpt. ≈17.1–17.4
98. 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
99. 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
00. Causal calculus: * Poole, Mackworth & Goebel 1998, pp. 335–337
01. Representing knowledge about knowledge: Belief calculus, modal logics: * Russell & Norvig 2003, pp. 341–344, * Poole, Mackworth & Goebel 1998, pp. 275–277

02. Sikos, Leslie F. (June 2017). *Description Logics in Multimedia Reasoning* (https://www.springer.com/us/book/9783319540658). Cham: Springer. doi:10.1007/978-3-319-54066-5 (https://doi.org/10.1007%2F978-3-319-54066-5). ISBN 978-3-319-54066-5. S2CID 3180114 (https://api.semanticscholar.org/CorpusID:3180114). Archived (https://web.archive.org/web/20170829120912/https://www.springer.com/us/book/9783319540658) from the original on 29 August 2017.
03. Ontology: * Russell & Norvig 2003, pp. 320–328
04. Smoliar, Stephen W.; Zhang, HongJiang (1994). "Content based video indexing and retrieval". *IEEE Multimedia*. **1** (2): 62–72. doi:10.1109/93.311653 (https://doi.org/10.1109%2F93.311653). S2CID 32710913 (https://api.semanticscholar.org/CorpusID:32710913).
05. 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 (https://doi.org/10.1016%2Fj.imavis.2007.08.013).
06. Kuperman, G. J.; Reichley, R. M.; Bailey, T. C. (1 July 2006). "Using Commercial Knowledge Bases for Clinical Decision Support: Opportunities, Hurdles, and Recommendations" (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1513681). *Journal of the American Medical Informatics Association*. **13** (4): 369–371. doi:10.1197/jamia.M2055 (https://doi.org/10.1197%2Fjamia.M2055). PMC 1513681 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1513681). PMID 16622160 (https://pubmed.ncbi.nlm.nih.gov/16622160).
07. 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 (https://doi.org/10.1017%2FS0269888905000408). S2CID 14987656 (https://api.semanticscholar.org/CorpusID:14987656).
08. 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.
09. Qualification problem: * McCarthy & Hayes 1969 * Russell & Norvig 2003 While McCarthy was primarily concerned with issues in the logical representation of actions, Russell & Norvig 2003 apply the term to the more general issue of default reasoning in the vast network of assumptions underlying all our commonsense knowledge.
10. Default reasoning and default logic, non-monotonic logics, circumscription, closed world assumption, abduction (Poole *et al.* places abduction under "default reasoning". Luger *et al.* places this under "uncertain reasoning"): * 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
11. Breadth of commonsense knowledge: * Russell & Norvig 2003, p. 21, * Crevier 1993, pp. 113–114, * Moravec 1988, p. 13, * Lenat & Guha 1989 (Introduction)
12. Dreyfus & Dreyfus 1986.
13. Gladwell 2005.
14. Expert knowledge as embodied intuition: * Dreyfus & Dreyfus 1986 (Hubert Dreyfus is a philosopher and critic of AI who was among the first to argue that most useful human knowledge was encoded sub-symbolically. See Dreyfus' critique of AI) * Gladwell 2005 (Gladwell's *Blink* is a popular introduction to sub-symbolic reasoning and knowledge.) * Hawkins & Blakeslee 2005 (Hawkins argues that sub-symbolic knowledge should be the primary focus of AI research.)
15. Planning: * ACM 1998, ~I.2.8, * Russell & Norvig 2003, pp. 375–459, * Poole, Mackworth & Goebel 1998, pp. 281–316, * Luger & Stubblefield 2004, pp. 314–329, * Nilsson 1998, chpt. 10.1–2, 22
16. Information value theory: * Russell & Norvig 2003, pp. 600–604
17. Classical planning: * Russell & Norvig 2003, pp. 375–430, * Poole, Mackworth & Goebel 1998, pp. 281–315, * Luger & Stubblefield 2004, pp. 314–329, * Nilsson 1998, chpt. 10.1–2, 22

18. Planning and acting in non-deterministic domains: conditional planning, execution monitoring, replanning and continuous planning: * Russell & Norvig 2003, pp. 430–449
19. Multi-agent planning and emergent behavior: * Russell & Norvig 2003, pp. 449–455
20. Turing 1950.
21. Solomonoff 1956.
22. Alan Turing discussed the centrality of learning as early as 1950, in his classic paper "Computing Machinery and Intelligence".[120] In 1956, at the original Dartmouth AI summer conference, Ray Solomonoff wrote a report on unsupervised probabilistic machine learning: "An Inductive Inference Machine".[121]
23. 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*."
24. Learning: * ACM 1998, I.2.6, * 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
25. Jordan, M. I.; Mitchell, T. M. (16 July 2015). "Machine learning: Trends, perspectives, and prospects". *Science*. **349** (6245): 255–260. Bibcode:2015Sci...349..255J (https://ui.adsabs.harvard.edu/abs/2015Sci...349..255J). doi:10.1126/science.aaa8415 (https://doi.org/10.1126%2Fscience.aaa8415). PMID 26185243 (https://pubmed.ncbi.nlm.nih.gov/26185243). S2CID 677218 (https://api.semanticscholar.org/CorpusID:677218).
26. Reinforcement learning: * Russell & Norvig 2003, pp. 763–788 * Luger & Stubblefield 2004, pp. 442–449
27. Natural language processing: * ACM 1998, I.2.7 * Russell & Norvig 2003, pp. 790–831 * Poole, Mackworth & Goebel 1998, pp. 91–104 * Luger & Stubblefield 2004, pp. 591–632
28. "Versatile question answering systems: seeing in synthesis" (https://www.academia.edu/2475776/Versatile_question_answering_systems_seeing_in_synthesis) Archived (https://web.archive.org/web/20160201125047/http://www.academia.edu/2475776/Versatile_question_answering_systems_seeing_in_synthesis) 1 February 2016 at the Wayback Machine, Mittal et al., IJIIDS, 5(2), 119–142, 2011
29. Applications of natural language processing, including information retrieval (i.e. text mining) and machine translation: * Russell & Norvig 2003, pp. 840–857, * Luger & Stubblefield 2004, pp. 623–630
30. 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 (https://doi.org/10.1109%2FMCI.2014.2307227). S2CID 206451986 (https://api.semanticscholar.org/CorpusID:206451986).
31. Vincent, James (7 November 2019). "OpenAI has published the text-generating AI it said was too dangerous to share" (https://www.theverge.com/2019/11/7/20953040/openai-text-generation-ai-gpt-2-full-model-release-1-5b-parameters). *The Verge*. Archived (https://web.archive.org/web/20200611054114/https://www.theverge.com/2019/11/7/20953040/openai-text-generation-ai-gpt-2-full-model-release-1-5b-parameters) from the original on 11 June 2020. Retrieved 11 June 2020.
32. Machine perception: * Russell & Norvig 2003, pp. 537–581, 863–898 * Nilsson 1998, ~chpt. 6
33. Speech recognition: * ACM 1998, ~I.2.7 * Russell & Norvig 2003, pp. 568–578
34. Object recognition: * Russell & Norvig 2003, pp. 885–892
35. Computer vision: * ACM 1998, I.2.10 * Russell & Norvig 2003, pp. 863–898 * Nilsson 1998, chpt. 6
36. Robotics: * ACM 1998, I.2.9, * Russell & Norvig 2003, pp. 901–942, * Poole, Mackworth & Goebel 1998, pp. 443–460
37. Moving and configuration space: * Russell & Norvig 2003, pp. 916–932

38. Tecuci 2012.
39. Robotic mapping (localization, etc): * Russell & Norvig 2003, pp. 908–915
40. 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 (https://arxiv.org/abs/1606.05830). Bibcode:2016arXiv160605830C (https://ui.adsabs.harvard.edu/abs/2016arXiv160605830C). doi:10.1109/TRO.2016.2624754 (https://doi.org/10.1109%2FTRO.2016.2624754). S2CID 2596787 (https://api.semanticscholar.org/CorpusID:2596787).
41. Moravec, Hans (1988). *Mind Children*. Harvard University Press. p. 15.
42. Chan, Szu Ping (15 November 2015). "This is what will happen when robots take over the world" (https://www.telegraph.co.uk/finance/economics/11994694/Heres-what-will-happen-when-robots-take-over-the-world.html). Archived (https://web.archive.org/web/20180424064029/https://www.telegraph.co.uk/finance/economics/11994694/Heres-what-will-happen-when-robots-take-over-the-world.html) from the original on 24 April 2018. Retrieved 23 April 2018.
43. "IKEA furniture and the limits of AI" (https://www.economist.com/news/leaders/21740735-humans-have-had-good-run-most-recent-breakthrough-robotics-it-clear). *The Economist*. 2018. Archived (https://web.archive.org/web/20180424014106/https://economist.com/news/leaders/21740735-humans-have-had-good-run-most-recent-breakthrough-robotics-it-clear) from the original on 24 April 2018. Retrieved 24 April 2018.
44. *Kismet*.
45. Thompson, Derek (2018). "What Jobs Will the Robots Take?" (https://www.theatlantic.com/business/archive/2014/01/what-jobs-will-the-robots-take/283239/). *The Atlantic*. Archived (https://web.archive.org/web/20180424202435/https://www.theatlantic.com/business/archive/2014/01/what-jobs-will-the-robots-take/283239/) from the original on 24 April 2018. Retrieved 24 April 2018.
46. Scassellati, Brian (2002). "Theory of mind for a humanoid robot". *Autonomous Robots*. **12** (1): 13–24. doi:10.1023/A:1013298507114 (https://doi.org/10.1023%2FA%3A1013298507114). S2CID 1979315 (https://api.semanticscholar.org/CorpusID:1979315).
47. Cao, Yongcan; Yu, Wenwu; Ren, Wei; Chen, Guanrong (February 2013). "An Overview of Recent Progress in the Study of Distributed Multi-Agent Coordination". *IEEE Transactions on Industrial Informatics*. **9** (1): 427–438. arXiv:1207.3231 (https://arxiv.org/abs/1207.3231). doi:10.1109/TII.2012.2219061 (https://doi.org/10.1109%2FTII.2012.2219061). S2CID 9588126 (https://api.semanticscholar.org/CorpusID:9588126).
48. Thro 1993.
49. Edelson 1991.
50. Tao & Tan 2005.
51. 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 (https://doi.org/10.1016%2Fj.inffus.2017.02.003). hdl:1893/25490 (https://hdl.handle.net/1893%2F25490).
52. Emotion and affective computing: * Minsky 2006
53. Waddell, Kaveh (2018). "Chatbots Have Entered the Uncanny Valley" (https://www.theatlantic.com/technology/archive/2017/04/uncanny-valley-digital-assistants/523806/). *The Atlantic*. Archived (https://web.archive.org/web/20180424202350/https://www.theatlantic.com/technology/archive/2017/04/uncanny-valley-digital-assistants/523806/) from the original on 24 April 2018. Retrieved 24 April 2018.
54. Pennachin, C.; Goertzel, B. (2007). *Contemporary Approaches to Artificial General Intelligence*. *Artificial General Intelligence. Cognitive Technologies*. Cognitive Technologies. Berlin, Heidelberg: Springer. doi:10.1007/978-3-540-68677-4_1 (https://doi.org/10.1007%2F978-3-540-68677-4_1). ISBN 978-3-540-23733-4.

55. Roberts, Jacob (2016). "Thinking Machines: The Search for Artificial Intelligence" (https://web.archive.org/web/20180819152455/https://www.sciencehistory.org/distillations/magazine/thinking-machines-the-search-for-artificial-intelligence). *Distillations*. Vol. 2 no. 2. pp. 14–23. Archived from the original (https://www.sciencehistory.org/distillations/magazine/thinking-machines-the-search-for-artificial-intelligence) on 19 August 2018. Retrieved 20 March 2018.
56. "The superhero of artificial intelligence: can this genius keep it in check?" (https://www.theguardian.com/technology/2016/feb/16/demis-hassabis-artificial-intelligence-deepmind-alphago). *the Guardian*. 16 February 2016. Archived (https://web.archive.org/web/20180423220101/https://www.theguardian.com/technology/2016/feb/16/demis-hassabis-artificial-intelligence-deepmind-alphago) from the original on 23 April 2018. Retrieved 26 April 2018.
57. Mnih, Volodymyr; Kavukcuoglu, Koray; Silver, David; Rusu, Andrei A.; Veness, Joel; Bellemare, Marc G.; Graves, Alex; Riedmiller, Martin; Fidjeland, Andreas K.; Ostrovski, Georg; Petersen, Stig; Beattie, Charles; Sadik, Amir; Antonoglou, Ioannis; King, Helen; Kumaran, Dharshan; Wierstra, Daan; Legg, Shane; Hassabis, Demis (26 February 2015). "Human-level control through deep reinforcement learning". *Nature*. **518** (7540): 529–533. Bibcode:2015Natur.518..529M (https://ui.adsabs.harvard.edu/abs/2015Natur.518..529M). doi:10.1038/nature14236 (https://doi.org/10.1038%2Fnature14236). PMID 25719670 (https://pubmed.ncbi.nlm.nih.gov/25719670). S2CID 205242740 (https://api.semanticscholar.org/CorpusID:205242740).
58. Sample, Ian (14 March 2017). "Google's DeepMind makes AI program that can learn like a human" (https://www.theguardian.com/global/2017/mar/14/googles-deepmind-makes-ai-program-that-can-learn-like-a-human). *the Guardian*. Archived (https://web.archive.org/web/20180426212908/https://www.theguardian.com/global/2017/mar/14/googles-deepmind-makes-ai-program-that-can-learn-like-a-human) from the original on 26 April 2018. Retrieved 26 April 2018.
59. "From not working to neural networking" (https://www.economist.com/news/special-report/21700756-artificial-intelligence-boom-based-old-idea-modern-twist-not). *The Economist*. 2016. Archived (https://web.archive.org/web/20161231203934/https://www.economist.com/news/special-report/21700756-artificial-intelligence-boom-based-old-idea-modern-twist-not) from the original on 31 December 2016. Retrieved 26 April 2018.
60. Domingos 2015.
61. Artificial brain arguments: AI requires a simulation of the operation of the human brain * Russell & Norvig 2003, p. 957 * Crevier 1993, pp. 271 and 279 A few of the people who make some form of the argument: * Moravec 1988 * Kurzweil 2005, p. 262 * Hawkins & Blakeslee 2005 The most extreme form of this argument (the brain replacement scenario) was put forward by Clark Glymour in the mid-1970s and was touched on by Zenon Pylyshyn and John Searle in 1980.
62. 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 (https://doi.org/10.1016%2Fj.neucom.2010.08.012).
63. Nilsson 1983, p. 10.
64. Nils Nilsson writes: "Simply put, there is wide disagreement in the field about what AI is all about."[163]
65. AI's immediate precursors: * McCorduck 2004, pp. 51–107 * Crevier 1993, pp. 27–32 * Russell & Norvig 2003, pp. 15, 940 * Moravec 1988, p. 3
66. Haugeland 1985, pp. 112–117
67. The most dramatic case of sub-symbolic AI being pushed into the background was the devastating critique of perceptrons by Marvin Minsky and Seymour Papert in 1969. See History of AI, AI winter, or Frank Rosenblatt.
68. Cognitive simulation, Newell and Simon, AI at CMU (then called Carnegie Tech): * McCorduck 2004, pp. 139–179, 245–250, 322–323 (EPAM) * Crevier 1993, pp. 145–149
69. Soar (history): * McCorduck 2004, pp. 450–451 * Crevier 1993, pp. 258–263

70. McCarthy and AI research at SAIL and SRI International: * McCorduck 2004, pp. 251–259 * Crevier 1993
71. AI research at Edinburgh and in France, birth of Prolog: * Crevier 1993, pp. 193–196 * Howe 1994
72. AI at MIT under Marvin Minsky in the 1960s : * McCorduck 2004, pp. 259–305 * Crevier 1993, pp. 83–102, 163–176 * Russell & Norvig 2003, p. 19
73. Cyc: * McCorduck 2004, p. 489, who calls it "a determinedly scruffy enterprise" * Crevier 1993, pp. 239–243 * Russell & Norvig 2003, p. 363−365 * Lenat & Guha 1989
74. Knowledge revolution: * McCorduck 2004, pp. 266–276, 298–300, 314, 421 * Russell & Norvig 2003, pp. 22–23
75. Frederick, Hayes-Roth; William, Murray; Leonard, Adelman. "Expert systems". *AccessScience*. doi:10.1036/1097-8542.248550 (https://doi.org/10.1036%2F1097-8542.248550).
76. Embodied approaches to AI: * McCorduck 2004, pp. 454–462 * Brooks 1990 * Moravec 1988
77. Weng et al. 2001.
78. Lungarella et al. 2003.
79. Asada et al. 2009.
80. Oudeyer 2010.
81. Revival of connectionism: * Crevier 1993, pp. 214–215 * Russell & Norvig 2003, p. 25
82. Computational intelligence * IEEE Computational Intelligence Society (http://www.ieee-cis.org/) Archived (https://web.archive.org/web/20080509191840/http://www.ieee-cis.org/) 9 May 2008 at the Wayback Machine
83. Hutson, Matthew (16 February 2018). "Artificial intelligence faces reproducibility crisis" (http://science.sciencemag.org/content/359/6377/725). *Science*. pp. 725–726. Bibcode:2018Sci...359..725H (https://ui.adsabs.harvard.edu/abs/2018Sci...359..725H). doi:10.1126/science.359.6377.725 (https://doi.org/10.1126%2Fscience.359.6377.725). Archived (https://web.archive.org/web/20180429092157/http://science.sciencemag.org/content/359/6377/725) from the original on 29 April 2018. Retrieved 28 April 2018.
84. Norvig 2012.
85. Langley 2011.
86. Katz 2012.
87. The intelligent agent paradigm: * Russell & Norvig 2003, pp. 27, 32–58, 968–972 * Poole, Mackworth & Goebel 1998, pp. 7–21 * Luger & Stubblefield 2004, pp. 235–240 * Hutter 2005, pp. 125–126 The definition used in this article, in terms of goals, actions, perception and environment, is due to Russell & Norvig (2003). Other definitions also include knowledge and learning as additional criteria.
88. Agent architectures, hybrid intelligent systems: * Russell & Norvig (2003, pp. 27, 932, 970–972) * Nilsson (1998, chpt. 25)
89. Hierarchical control system: * Albus 2002
90. Lieto, Antonio; Lebiere, Christian; Oltramari, Alessandro (May 2018). "The knowledge level in cognitive architectures: Current limitations and possibile developments". *Cognitive Systems Research*. **48**: 39–55. doi:10.1016/j.cogsys.2017.05.001 (https://doi.org/10.1016%2Fj.cogsys.2017.05.001). hdl:2318/1665207 (https://hdl.handle.net/2318%2F1665207). S2CID 206868967 (https://api.semanticscholar.org/CorpusID:206868967).
91. Lieto, Antonio; Bhatt, Mehul; Oltramari, Alessandro; Vernon, David (May 2018). "The role of cognitive architectures in general artificial intelligence". *Cognitive Systems Research*. **48**: 1–3. doi:10.1016/j.cogsys.2017.08.003 (https://doi.org/10.1016%2Fj.cogsys.2017.08.003). hdl:2318/1665249 (https://hdl.handle.net/2318%2F1665249). S2CID 36189683 (https://api.semanticscholar.org/CorpusID:36189683).
92. Russell & Norvig 2009, p. 1.

93. *White Paper: On Artificial Intelligence - A European approach to excellence and trust* (https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf) (PDF). Brussels: European Commission. 2020. p. 1. Archived (https://web.archive.org/web/20200220173419/https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf) (PDF) from the original on 20 February 2020. Retrieved 20 February 2020.
94. *CNN* 2006.
95. Using AI to predict flight delays (https://ishti.org/2018/11/19/using-artificial-intelligence-to-predict-flight-delays/) Archived (https://web.archive.org/web/20181120095602/https://ishti.org/2018/11/19/using-artificial-intelligence-to-predict-flight-delays/) 20 November 2018 at the Wayback Machine, Ishti.org.
96. N. Aletras; D. Tsarapatsanis; D. Preotiuc-Pietro; V. Lampos (2016). "Predicting judicial decisions of the European Court of Human Rights: a Natural Language Processing perspective" (https://doi.org/10.7717%2Fpeerj-cs.93). *PeerJ Computer Science*. **2**: e93. doi:10.7717/peerj-cs.93 (https://doi.org/10.7717%2Fpeerj-cs.93).
97. "The Economist Explains: Why firms are piling into artificial intelligence" (https://www.economist.com/blogs/economist-explains/2016/04/economist-explains). *The Economist*. 31 March 2016. Archived (https://web.archive.org/web/20160508010311/http://www.economist.com/blogs/economist-explains/2016/04/economist-explains) from the original on 8 May 2016. Retrieved 19 May 2016.
98. Lohr, Steve (28 February 2016). "The Promise of Artificial Intelligence Unfolds in Small Steps" (https://www.nytimes.com/2016/02/29/technology/the-promise-of-artificial-intelligence-unfolds-in-small-steps.html). *The New York Times*. Archived (https://web.archive.org/web/20160229171843/http://www.nytimes.com/2016/02/29/technology/the-promise-of-artificial-intelligence-unfolds-in-small-steps.html) from the original on 29 February 2016. Retrieved 29 February 2016.
99. Frangoul, Anmar (14 June 2019). "A Californian business is using A.I. to change the way we think about energy storage" (https://www.cnbc.com/2019/06/14/the-business-using-ai-to-change-how-we-think-about-energy-storage.html). *CNBC*. Archived (https://web.archive.org/web/20200725044735/https://www.cnbc.com/2019/06/14/the-business-using-ai-to-change-how-we-think-about-energy-storage.html) from the original on 25 July 2020. Retrieved 5 November 2019.
00. Wakefield, Jane (15 June 2016). "Social media 'outstrips TV' as news source for young people" (https://www.bbc.co.uk/news/uk-36528256). *BBC News*. Archived (https://web.archive.org/web/20160624000744/http://www.bbc.co.uk/news/uk-36528256) from the original on 24 June 2016.
01. Smith, Mark (22 July 2016). "So you think you chose to read this article?" (https://www.bbc.co.uk/news/business-36837824). *BBC News*. Archived (https://web.archive.org/web/20160725205007/http://www.bbc.co.uk/news/business-36837824) from the original on 25 July 2016.
02. Brown, Eileen. "Half of Americans do not believe deepfake news could target them online" (https://www.zdnet.com/article/half-of-americans-do-not-believe-deepfake-news-could-target-them-online/). *ZDNet*. Archived (https://web.archive.org/web/20191106035012/https://www.zdnet.com/article/half-of-americans-do-not-believe-deepfake-news-could-target-them-online/) from the original on 6 November 2019. Retrieved 3 December 2019.
03. The Turing test:
Turing's original publication: * Turing 1950 Historical influence and philosophical implications: * Haugeland 1985, pp. 6–9 * Crevier 1993, p. 24 * McCorduck 2004, pp. 70–71 * Russell & Norvig 2003, pp. 2–3 and 948
04. Dartmouth proposal: * McCarthy et al. 1955 (the original proposal) * Crevier 1993, p. 49 (historical significance)
05. The physical symbol systems hypothesis: * Newell & Simon 1976, p. 116 * McCorduck 2004, p. 153 * Russell & Norvig 2003, p. 18
06. Dreyfus 1992, p. 156.

07. Dreyfus criticized the necessary condition of the physical symbol system hypothesis, which he called the "psychological assumption": "The mind can be viewed as a device operating on bits of information according to formal rules."[206]
08. Dreyfus' critique of artificial intelligence: * Dreyfus 1972, Dreyfus & Dreyfus 1986 * Crevier 1993, pp. 120–132 * McCorduck 2004, pp. 211–239 * Russell & Norvig 2003, pp. 950–952,
09. Gödel 1951: in this lecture, Kurt Gödel uses the incompleteness theorem to arrive at the following disjunction: (a) the human mind is not a consistent finite machine, or (b) there exist Diophantine equations for which it cannot decide whether solutions exist. Gödel finds (b) implausible, and thus seems to have believed the human mind was not equivalent to a finite machine, i.e., its power exceeded that of any finite machine. He recognized that this was only a conjecture, since one could never disprove (b). Yet he considered the disjunctive conclusion to be a "certain fact".
10. The Mathematical Objection: * Russell & Norvig 2003, p. 949 * McCorduck 2004, pp. 448–449 Making the Mathematical Objection: * Lucas 1961 * Penrose 1989 Refuting Mathematical Objection: * Turing 1950 under "(2) The Mathematical Objection" * Hofstadter 1979 Background: * Gödel 1931, Church 1936, Kleene 1935, Turing 1937
11. Graham Oppy (20 January 2015). "Gödel's Incompleteness Theorems" (http://plato.stanford.edu/entries/goedel-incompleteness/#GdeArgAgaMec). *Stanford Encyclopedia of Philosophy*. Archived (https://web.archive.org/web/20160422031755/http://plato.stanford.edu/entries/goedel-incompleteness/#GdeArgAgaMec) from the original on 22 April 2016. Retrieved 27 April 2016. "These Gödelian anti-mechanist arguments are, however, problematic, and there is wide consensus that they fail."
12. Stuart J. Russell; Peter Norvig (2010). "26.1.2: Philosophical Foundations/Weak AI: Can Machines Act Intelligently?/The mathematical objection". *Artificial Intelligence: A Modern Approach* (3rd ed.). Upper Saddle River, NJ: Prentice Hall. ISBN 978-0-13-604259-4. "even if we grant that computers have limitations on what they can prove, there is no evidence that humans are immune from those limitations."
13. Mark Colyvan. An introduction to the philosophy of mathematics. Cambridge University Press, 2012. From 2.2.2, 'Philosophical significance of Gödel's incompleteness results': "The accepted wisdom (with which I concur) is that the Lucas-Penrose arguments fail."
14. Iphofen, Ron; Kritikos, Mihalis (3 January 2019). "Regulating artificial intelligence and robotics: ethics by design in a digital society". *Contemporary Social Science*: 1–15. doi:10.1080/21582041.2018.1563803 (https://doi.org/10.1080%2F21582041.2018.1563803). ISSN 2158-2041 (https://www.worldcat.org/issn/2158-2041).
15. "Ethical AI Learns Human Rights Framework" (https://www.voanews.com/episode/ethical-ai-learns-human-rights-framework-4087171). *Voice of America*. Archived (https://web.archive.org/web/20191111000034/https://www.voanews.com/episode/ethical-ai-learns-human-rights-framework-4087171) from the original on 11 November 2019. Retrieved 10 November 2019.
16. Crevier 1993, pp. 132–144.
17. In the early 1970s, Kenneth Colby presented a version of Weizenbaum's ELIZA known as DOCTOR which he promoted as a serious therapeutic tool.[216]
18. Joseph Weizenbaum's critique of AI: * Weizenbaum 1976 * Crevier 1993, pp. 132–144 * McCorduck 2004, pp. 356–373 * Russell & Norvig 2003, p. 961 Weizenbaum (the AI researcher who developed the first chatterbot program, ELIZA) argued in 1976 that the misuse of artificial intelligence has the potential to devalue human life.
19. Wendell Wallach (2010). *Moral Machines*, Oxford University Press.
20. Wallach, pp 37–54.
21. Wallach, pp 55–73.
22. Wallach, Introduction chapter.
23. Michael Anderson and Susan Leigh Anderson (2011), Machine Ethics, Cambridge University Press.

24. "Machine Ethics" (https://web.archive.org/web/20141129044821/http://www.aaai.org/Library/Symposia/Fall/fs05-06). *aaai.org*. Archived from the original (http://www.aaai.org/Library/Symposia/Fall/fs05-06) on 29 November 2014.
25. Rubin, Charles (Spring 2003). "Artificial Intelligence and Human Nature" (https://web.archive.org/web/20120611115223/http://www.thenewatlantis.com/publications/artificial-intelligence-and-human-nature). *The New Atlantis*. **1**: 88–100. Archived from the original (http://www.thenewatlantis.com/publications/artificial-intelligence-and-human-nature) on 11 June 2012.
26. Brooks, Rodney (10 November 2014). "artificial intelligence is a tool, not a threat" (https://web.archive.org/web/20141112130954/http://www.rethinkrobotics.com/artificial-intelligence-tool-threat/). Archived from the original (http://www.rethinkrobotics.com/artificial-intelligence-tool-threat/) on 12 November 2014.
27. "Stephen Hawking, Elon Musk, and Bill Gates Warn About Artificial Intelligence" (https://observer.com/2015/08/stephen-hawking-elon-musk-and-bill-gates-warn-about-artificial-intelligence/). *Observer*. 19 August 2015. Archived (https://web.archive.org/web/20151030053323/http://observer.com/2015/08/stephen-hawking-elon-musk-and-bill-gates-warn-about-artificial-intelligence/) from the original on 30 October 2015. Retrieved 30 October 2015.
28. Chalmers, David (1995). "Facing up to the problem of consciousness" (http://www.imprint.co.uk/chalmers.html). *Journal of Consciousness Studies*. **2** (3): 200–219. Archived (https://web.archive.org/web/20050308163649/http://www.imprint.co.uk/chalmers.html) from the original on 8 March 2005. Retrieved 11 October 2018. See also this link (http://consc.net/papers/facing.html) Archived (https://web.archive.org/web/20110408224800/http://consc.net/papers/facing.html) 8 April 2011 at the Wayback Machine
29. Horst, Steven, (2005) "The Computational Theory of Mind" (http://plato.stanford.edu/entries/computational-mind/) Archived (https://web.archive.org/web/20180911043901/https://plato.stanford.edu/entries/computational-mind/) 11 September 2018 at the Wayback Machine in *The Stanford Encyclopedia of Philosophy*
30. Searle 1980, p. 1.
31. This version is from Searle (1999), and is also quoted in Dennett 1991, p. 435. 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." [230] Strong AI is defined similarly by Russell & Norvig (2003, p. 947): "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."
32. 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
33. Robot rights: * Russell & Norvig 2003, p. 964 * *BBC News* 2006 Prematurity of: * Henderson 2007 In fiction: * McCorduck (2004, pp. 190–25) discusses *Frankenstein* and identifies the key ethical issues as scientific hubris and the suffering of the monster, i.e. robot rights.
34. Evans, Woody (2015). "Posthuman Rights: Dimensions of Transhuman Worlds" (https://doi.org/10.5209%2Frev_TK.2015.v12.n2.49072). *Teknokultura*. **12** (2). doi:10.5209/rev_TK.2015.v12.n2.49072 (https://doi.org/10.5209%2Frev_TK.2015.v12.n2.49072).
35. maschafilm. "Content: Plug & Pray Film – Artificial Intelligence – Robots -" (http://www.plugandpray-film.de/en/content.html). *plugandpray-film.de*. Archived (https://web.archive.org/web/20160212040134/http://www.plugandpray-film.de/en/content.html) from the original on 12 February 2016.
36. Omohundro, Steve (2008). *The Nature of Self-Improving Artificial Intelligence*. presented and distributed at the 2007 Singularity Summit, San Francisco, CA.
37. Technological singularity: * Vinge 1993 * Kurzweil 2005 * Russell & Norvig 2003, p. 963
38. Transhumanism: * Moravec 1988 * Kurzweil 2005 * Russell & Norvig 2003, p. 963

39. AI as evolution: * Edward Fredkin is quoted in McCorduck (2004, p. 401). * Butler 1863 * Dyson 1998
40. "Robots and Artificial Intelligence" (http://www.igmchicago.org/surveys/robots-and-artificial-intelligence). *www.igmchicago.org*. Archived (https://web.archive.org/web/20190501114826/http://www.igmchicago.org/surveys/robots-and-artificial-intelligence) from the original on 1 May 2019. Retrieved 3 July 2019.
41. E McGaughey, 'Will Robots Automate Your Job Away? Full Employment, Basic Income, and Economic Democracy' (2018) SSRN, part 2(3) (https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3044448) Archived (https://web.archive.org/web/20180524201340/https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3044448) 24 May 2018 at the Wayback Machine
42. "Automation and anxiety" (https://www.economist.com/news/special-report/21700758-will-smarter-machines-cause-mass-unemployment-automation-and-anxiety). *The Economist*. 9 May 2015. Archived (https://web.archive.org/web/20180112214621/https://www.economist.com/news/special-report/21700758-will-smarter-machines-cause-mass-unemployment-automation-and-anxiety) from the original on 12 January 2018. Retrieved 13 January 2018.
43. Lohr, Steve (2017). "Robots Will Take Jobs, but Not as Fast as Some Fear, New Report Says" (https://www.nytimes.com/2017/01/12/technology/robots-will-take-jobs-but-not-as-fast-as-some-fear-new-report-says.html). *The New York Times*. Archived (https://web.archive.org/web/20180114073704/https://www.nytimes.com/2017/01/12/technology/robots-will-take-jobs-but-not-as-fast-as-some-fear-new-report-says.html) from the original on 14 January 2018. Retrieved 13 January 2018.
44. 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 (https://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.395.416). doi:10.1016/j.techfore.2016.08.019 (https://doi.org/10.1016%2Fj.techfore.2016.08.019). ISSN 0040-1625 (https://www.worldcat.org/issn/0040-1625).
45. Arntz, Melanie, Terry Gregory, and Ulrich Zierahn. "The risk of automation for jobs in OECD countries: A comparative analysis." OECD Social, Employment, and Migration Working Papers 189 (2016). p. 33.
46. Mahdawi, Arwa (26 June 2017). "What jobs will still be around in 20 years? Read this to prepare your future" (https://www.theguardian.com/us-news/2017/jun/26/jobs-future-automation-robots-skills-creative-health). *The Guardian*. Archived (https://web.archive.org/web/20180114021804/https://www.theguardian.com/us-news/2017/jun/26/jobs-future-automation-robots-skills-creative-health) from the original on 14 January 2018. Retrieved 13 January 2018.
47. Simon, Matt (1 April 2019). "Andrew Yang's Presidential Bid Is So Very 21st Century" (https://www.wired.com/story/andrew-yangs-presidential-bid-is-so-very-21st-century/). *Wired*. Archived (https://web.archive.org/web/20190624151143/https://www.wired.com/story/andrew-yangs-presidential-bid-is-so-very-21st-century/) from the original on 24 June 2019. Retrieved 2 May 2019 – via www.wired.com.
48. "Five experts share what scares them the most about AI" (https://futurism.com/artificial-intelligence-experts-fear/amp). 5 September 2018. Archived (https://web.archive.org/web/20191208094101/https://futurism.com/artificial-intelligence-experts-fear/amp) from the original on 8 December 2019. Retrieved 8 December 2019.
49. Russel, Stuart., Daniel Dewey, and Max Tegmark. Research Priorities for Robust and Beneficial Artificial Intelligence. AI Magazine 36:4 (2015). 8 December 2016.
50. "Commentary: Bad news. Artificial intelligence is biased" (https://www.channelnewsasia.com/news/commentary/artificial-intelligence-big-data-bias-hiring-loans-key-challenge-11097374). *CNA*. 12 January 2019. Archived (https://web.archive.org/web/20190112104421/https://www.channelnewsasia.com/news/commentary/artificial-intelligence-big-data-bias-hiring-loans-key-challenge-11097374) from the original on 12 January 2019. Retrieved 19 June 2020.

51. Jeff Larson, Julia Angwin (23 May 2016). "How We Analyzed the COMPAS Recidivism Algorithm" (https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm). *ProPublica*. Archived (https://web.archive.org/web/20190429190950/https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm) from the original on 29 April 2019. Retrieved 19 June 2020.
52. Rawlinson, Kevin (29 January 2015). "Microsoft's Bill Gates insists AI is a threat" (https://www.bbc.co.uk/news/31047780). *BBC News*. Archived (https://web.archive.org/web/20150129183607/http://www.bbc.co.uk/news/31047780) from the original on 29 January 2015. Retrieved 30 January 2015.
53. Holley, Peter (28 January 2015). "Bill Gates on dangers of artificial intelligence: 'I don't understand why some people are not concerned' " (https://www.washingtonpost.com/news/the-switch/wp/2015/01/28/bill-gates-on-dangers-of-artificial-intelligence-dont-understand-why-some-people-are-not-concerned/). *The Washington Post*. ISSN 0190-8286 (https://www.worldcat.org/issn/0190-8286). Archived (https://web.archive.org/web/20151030054330/https://www.washingtonpost.com/news/the-switch/wp/2015/01/28/bill-gates-on-dangers-of-artificial-intelligence-dont-understand-why-some-people-are-not-concerned/) from the original on 30 October 2015. Retrieved 30 October 2015.
54. Gibbs, Samuel (27 October 2014). "Elon Musk: artificial intelligence is our biggest existential threat" (https://www.theguardian.com/technology/2014/oct/27/elon-musk-artificial-intelligence-ai-biggest-existential-threat). *The Guardian*. Archived (https://web.archive.org/web/20151030054330/http://www.theguardian.com/technology/2014/oct/27/elon-musk-artificial-intelligence-ai-biggest-existential-threat) from the original on 30 October 2015. Retrieved 30 October 2015.
55. Cellan-Jones, Rory (2 December 2014). "Stephen Hawking warns artificial intelligence could end mankind" (https://www.bbc.com/news/technology-30290540). *BBC News*. Archived (https://web.archive.org/web/20151030054329/http://www.bbc.com/news/technology-30290540) from the original on 30 October 2015. Retrieved 30 October 2015.
56. Bostrom, Nick (2015). "What happens when our computers get smarter than we are?" (https://www.ted.com/talks/nick_bostrom_what_happens_when_our_computers_get_smarter_than_we_are/transcript). TED (conference). Archived (https://web.archive.org/web/20200725005719/https://www.ted.com/talks/nick_bostrom_what_happens_when_our_computers_get_smarter_than_we_are/transcript) from the original on 25 July 2020. Retrieved 30 January 2020.
57. 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 (https://www.worldcat.org/oclc/1083694322).
58. Post, Washington. "Tech titans like Elon Musk are spending $1 billion to save you from terminators" (https://www.chicagotribune.com/bluesky/technology/ct-tech-titans-against-terminators-20151214-story.html). Archived (https://web.archive.org/web/20160607121118/http://www.chicagotribune.com/bluesky/technology/ct-tech-titans-against-terminators-20151214-story.html) from the original on 7 June 2016.
59. Müller, Vincent C.; Bostrom, Nick (2014). "Future Progress in Artificial Intelligence: A Poll Among Experts" (http://www.sophia.de/pdf/2014_PT-AI_polls.pdf) (PDF). *AI Matters*. **1** (1): 9–11. doi:10.1145/2639475.2639478 (https://doi.org/10.1145%2F2639475.2639478). S2CID 8510016 (https://api.semanticscholar.org/CorpusID:8510016). Archived (https://web.archive.org/web/20160115114604/http://www.sophia.de/pdf/2014_PT-AI_polls.pdf) (PDF) from the original on 15 January 2016.
60. "Oracle CEO Mark Hurd sees no reason to fear ERP AI" (https://searcherp.techtarget.com/news/252460208/Oracle-CEO-Mark-Hurd-sees-no-reason-to-fear-ERP-AI). *SearchERP*. Archived (https://web.archive.org/web/20190506173749/https://searcherp.techtarget.com/news/252460208/Oracle-CEO-Mark-Hurd-sees-no-reason-to-fear-ERP-AI) from the original on 6 May 2019. Retrieved 6 May 2019.

61. "Mark Zuckerberg responds to Elon Musk's paranoia about AI: 'AI is going to... help keep our communities safe.' " (https://www.businessinsider.com/mark-zuckerberg-shares-thoughts-elon-musks-ai-2018-5). *Business Insider*. 25 May 2018. Archived (https://web.archive.org/web/20190506173756/https://www.businessinsider.com/mark-zuckerberg-shares-thoughts-elon-musks-ai-2018-5) from the original on 6 May 2019. Retrieved 6 May 2019.
62. "The mysterious artificial intelligence company Elon Musk invested in is developing game-changing smart computers" (http://www.techinsider.io/mysterious-artificial-intelligence-company-elon-musk-investment-2015-10). *Tech Insider*. Archived (https://web.archive.org/web/20151030165333/http://www.techinsider.io/mysterious-artificial-intelligence-company-elon-musk-investment-2015-10) from the original on 30 October 2015. Retrieved 30 October 2015.
63. Clark, Jack. "Musk-Backed Group Probes Risks Behind Artificial Intelligence" (https://www.bloomberg.com/news/articles/2015-07-01/musk-backed-group-probes-risks-behind-artificial-intelligence). *Bloomberg.com*. Archived (https://web.archive.org/web/20151030202356/http://www.bloomberg.com/news/articles/2015-07-01/musk-backed-group-probes-risks-behind-artificial-intelligence) from the original on 30 October 2015. Retrieved 30 October 2015.
64. "Elon Musk Is Donating $10M Of His Own Money To Artificial Intelligence Research" (http://www.fastcompany.com/3041007/fast-feed/elon-musk-is-donating-10m-of-his-own-money-to-artificial-intelligence-research). *Fast Company*. 15 January 2015. Archived (https://web.archive.org/web/20151030202356/http://www.fastcompany.com/3041007/fast-feed/elon-musk-is-donating-10m-of-his-own-money-to-artificial-intelligence-research) from the original on 30 October 2015. Retrieved 30 October 2015.
65. "Is artificial intelligence really an existential threat to humanity?" (http://thebulletin.org/artificial-intelligence-really-existential-threat-humanity8577). *Bulletin of the Atomic Scientists*. 9 August 2015. Archived (https://web.archive.org/web/20151030054330/http://thebulletin.org/artificial-intelligence-really-existential-threat-humanity8577) from the original on 30 October 2015. Retrieved 30 October 2015.
66. "The case against killer robots, from a guy actually working on artificial intelligence" (http://fusion.net/story/54583/the-case-against-killer-robots-from-a-guy-actually-building-ai/). *Fusion.net*. Archived (https://web.archive.org/web/20160204175716/http://fusion.net/story/54583/the-case-against-killer-robots-from-a-guy-actually-building-ai/) from the original on 4 February 2016. Retrieved 31 January 2016.
67. "Will artificial intelligence destroy humanity? Here are 5 reasons not to worry" (https://www.vox.com/2014/8/22/6043635/5-reasons-we-shouldnt-worry-about-super-intelligent-computers-taking). *Vox*. 22 August 2014. Archived (https://web.archive.org/web/20151030092203/http://www.vox.com/2014/8/22/6043635/5-reasons-we-shouldnt-worry-about-super-intelligent-computers-taking) from the original on 30 October 2015. Retrieved 30 October 2015.
68. Berryhill, Jamie; Heang, Kévin Kok; Clogher, Rob; McBride, Keegan (2019). *Hello, World: Artificial Intelligence and its Use in the Public Sector* (https://oecd-opsi.org/wp-content/uploads/2019/11/AI-Report-Online.pdf) (PDF). Paris: OECD Observatory of Public Sector Innovation. Archived (https://web.archive.org/web/20191220021331/https://oecd-opsi.org/wp-content/uploads/2019/11/AI-Report-Online.pdf) (PDF) from the original on 20 December 2019. Retrieved 9 August 2020.
69. *Research handbook on the law of artificial intelligence*. Barfield, Woodrow,, Pagallo, Ugo. Cheltenham, UK. 2018. ISBN 978-1-78643-904-8. OCLC 1039480085 (https://www.worldcat.org/oclc/1039480085).
70. Law Library of Congress (U.S.). Global Legal Research Directorate, issuing body. *Regulation of artificial intelligence in selected jurisdictions*. LCCN 2019668143 (https://lccn.loc.gov/2019668143). OCLC 1110727808 (https://www.worldcat.org/oclc/1110727808).

71. Wirtz, Bernd W.; Weyerer, Jan C.; Geyer, Carolin (24 July 2018). "Artificial Intelligence and the Public Sector—Applications and Challenges" (https://zenodo.org/record/3569435). *International Journal of Public Administration*. **42** (7): 596–615. doi:10.1080/01900692.2018.1498103 (https://doi.org/10.1080%2F01900692.2018.1498103). ISSN 0190-0692 (https://www.worldcat.org/issn/0190-0692). S2CID 158829602 (https://api.semanticscholar.org/CorpusID:158829602). Archived (https://web.archive.org/web/20200818131415/https://zenodo.org/record/3569435) from the original on 18 August 2020. Retrieved 22 August 2020.
72. Buiten, Miriam C (2019). "Towards Intelligent Regulation of Artificial Intelligence" (https://doi.org/10.1017%2Ferr.2019.8). *European Journal of Risk Regulation*. **10** (1): 41–59. doi:10.1017/err.2019.8 (https://doi.org/10.1017%2Ferr.2019.8). ISSN 1867-299X (https://www.worldcat.org/issn/1867-299X).
73. Sotala, Kaj; Yampolskiy, Roman V (19 December 2014). "Responses to catastrophic AGI risk: a survey" (https://doi.org/10.1088%2F0031-8949%2F90%2F1%2F018001). *Physica Scripta*. **90** (1): 018001. doi:10.1088/0031-8949/90/1/018001 (https://doi.org/10.1088%2F0031-8949%2F90%2F1%2F018001). ISSN 0031-8949 (https://www.worldcat.org/issn/0031-8949).
74. Buttazzo, G. (July 2001). "Artificial consciousness: Utopia or real possibility?". *Computer*. **34** (7): 24–30. doi:10.1109/2.933500 (https://doi.org/10.1109%2F2.933500).
75. Anderson, Susan Leigh. "Asimov's "three laws of robotics" and machine metaethics." AI & Society 22.4 (2008): 477–493.
76. 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 (https://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.85.8904). doi:10.1007/s10676-007-9138-2 (https://doi.org/10.1007%2Fs10676-007-9138-2). S2CID 37272949 (https://api.semanticscholar.org/CorpusID:37272949).
77. 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 (https://www.jstor.org/stable/4240644).

AI textbooks

- Hutter, Marcus (2005). *Universal Artificial Intelligence*. Berlin: Springer. ISBN 978-3-540-22139-5.
- Jackson, Philip (1985). *Introduction to Artificial Intelligence* (https://archive.org/details/introductiontoar1985jack) (2nd ed.). Dover. ISBN 978-0-486-24864-6. Archived (https://web.archive.org/web/20200726131713/https://archive.org/details/introductiontoar1985jack) from the original on 26 July 2020. Retrieved 4 March 2020.
- Luger, George; Stubblefield, William (2004). *Artificial Intelligence: Structures and Strategies for Complex Problem Solving* (https://archive.org/details/artificialintell0000luge) (5th ed.). Benjamin/Cummings. ISBN 978-0-8053-4780-7. Archived (https://web.archive.org/web/20200726220613/https://archive.org/details/artificialintell0000luge) from the original on 26 July 2020. Retrieved 17 December 2019.
- Neapolitan, Richard; Jiang, Xia (2018). *Artificial Intelligence: With an Introduction to Machine Learning* (https://www.crcpress.com/Contemporary-Artificial-Intelligence-Second-Edition/Neapolitan-Jiang/p/book/9781138502383). Chapman & Hall/CRC. ISBN 978-1-138-50238-3. Archived (https://web.archive.org/web/20200822201555/https://www.routledge.com/Contemporary-Artificial-Intelligence-Second-Edition/Neapolitan-Jiang/p/book/9781138502383) from the original on 22 August 2020. Retrieved 3 January 2018.
- Nilsson, Nils (1998). *Artificial Intelligence: A New Synthesis* (https://archive.org/details/artificialintell0000nils). Morgan Kaufmann. ISBN 978-1-55860-467-4. Archived (https://web.archive.org/web/20200726131654/https://archive.org/details/artificialintell0000nils) from the original on 26 July 2020. Retrieved 18 November 2019.

- Russell, Stuart J.; Norvig, Peter (2003), *Artificial Intelligence: A Modern Approach* (http://aima.cs.berkeley.edu/) (2nd ed.), Upper Saddle River, New Jersey: Prentice Hall, ISBN 0-13-790395-2.
- 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; Goebel, Randy (1998). *Computational Intelligence: A Logical Approach* (https://archive.org/details/computationalint00pool). New York: Oxford University Press. ISBN 978-0-19-510270-3. Archived (https://web.archive.org/web/20200726131436/https://archive.org/details/computationalint00pool) from the original on 26 July 2020. Retrieved 22 August 2020.
- Winston, Patrick Henry (1984). *Artificial Intelligence* (https://archive.org/details/artificialintell00wins). Reading, MA: Addison-Wesley. ISBN 978-0-201-08259-3. Archived (https://web.archive.org/web/20200726131500/https://archive.org/details/artificialintell00wins) from the original on 26 July 2020. Retrieved 22 August 2020.
- Rich, Elaine (1983). *Artificial Intelligence* (https://archive.org/details/ine0000unse). McGraw-Hill. ISBN 978-0-07-052261-9. Archived (https://web.archive.org/web/20200726131632/https://archive.org/details/ine0000unse) from the original on 26 July 2020. Retrieved 17 December 2019.
- Bundy, Alan (1980). *Artificial Intelligence: An Introductory Course* (2nd ed.). Edinburgh University Press. ISBN 978-0-85224-410-4.
- Poole, David; Mackworth, Alan (2017). *Artificial Intelligence: Foundations of Computational Agents* (http://artint.info/index.html) (2nd ed.). Cambridge University Press. ISBN 978-1-107-19539-4. Archived (https://web.archive.org/web/20171207013855/http://artint.info/index.html) from the original on 7 December 2017. Retrieved 6 December 2017.

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* (http://www.pamelamc.com/html/machines_who_think.html) (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

- 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 (https://doi.org/10.1109%2Ftamd.2009.2021702). S2CID 10168773 (https://api.semanticscholar.org/CorpusID:10168773).
- "ACM Computing Classification System: Artificial intelligence" (https://web.archive.org/web/20071012025921/http://www.acm.org/class/1998/I.2.html). ACM. 1998. Archived from the original (http://www.acm.org/class/1998/I.2.html) on 12 October 2007. Retrieved 30 August 2007.
- Goodman, Joanna (2016). *Robots in Law: How Artificial Intelligence is Transforming Legal Services* (https://www.ark-group.com/product/robots-law-how-artificial-intelligence-transforming-legal-services) (1st ed.). Ark Group. ISBN 978-1-78358-264-8. Archived (https://web.archive.org/web/20161108052144/https://www.ark-group.com/product/robots-law-how-artificial-intelligence-transforming-legal-services) from the original on 8 November 2016. Retrieved 7 November 2016.

- Albus, J. S. (2002). "4-D/RCS: A Reference Model Architecture for Intelligent Unmanned Ground Vehicles" (https://web.archive.org/web/20040725051856/http://www.isd.mel.nist.gov/documents/albus/4DRCS.pdf) (PDF). In Gerhart, G.; Gunderson, R.; Shoemaker, C. (eds.). *Proceedings of the SPIE AeroSense Session on Unmanned Ground Vehicle Technology*. Unmanned Ground Vehicle Technology IV. **3693**. pp. 11–20. Bibcode:2002SPIE.4715..303A (https://ui.adsabs.harvard.edu/abs/2002SPIE.4715..303A). CiteSeerX 10.1.1.15.14 (https://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.15.14). doi:10.1117/12.474462 (https://doi.org/10.1117%2F12.474462). S2CID 63339739 (https://api.semanticscholar.org/CorpusID:63339739). Archived from the original (http://www.isd.mel.nist.gov/documents/albus/4DRCS.pdf) (PDF) on 25 July 2004.
- Aleksander, Igor (1995). *Artificial Neuroconsciousness: An Update* (https://web.archive.org/web/19970302014628/http://www.ee.ic.ac.uk/research/neural/publications/iwann.html). IWANN. Archived from the original (http://www.ee.ic.ac.uk/research/neural/publications/iwann.html) on 2 March 1997. BibTex (http://dblp.uni-trier.de/rec/bibtex/conf/iwann/Aleksander95) Archived (https://web.archive.org/web/19970302014628/http://www.ee.ic.ac.uk/research/neural/publications/iwann.html) 2 March 1997 at the Wayback Machine.
- Bach, Joscha (2008). "Seven Principles of Synthetic Intelligence" (https://books.google.com/books?id=a_ZR81Z25z0C&pg=PA63). In Wang, Pei; Goertzel, Ben; Franklin, Stan (eds.). *Artificial General Intelligence, 2008: Proceedings of the First AGI Conference*. IOS Press. pp. 63–74. ISBN 978-1-58603-833-5. Archived (https://web.archive.org/web/20160708030627/https://books.google.com/books?id=a_ZR81Z25z0C&pg=PA63) from the original on 8 July 2016. Retrieved 16 February 2016.
- "Robots could demand legal rights" (http://news.bbc.co.uk/2/hi/technology/6200005.stm). *BBC News*. 21 December 2006. Archived (https://web.archive.org/web/20191015042628/http://news.bbc.co.uk/2/hi/technology/6200005.stm) from the original on 15 October 2019. Retrieved 3 February 2011.
- Brooks, Rodney (1990). "Elephants Don't Play Chess" (http://people.csail.mit.edu/brooks/papers/elephants.pdf) (PDF). *Robotics and Autonomous Systems*. **6** (1–2): 3–15. CiteSeerX 10.1.1.588.7539 (https://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.588.7539). doi:10.1016/S0921-8890(05)80025-9 (https://doi.org/10.1016%2FS0921-8890%2805%2980025-9). Archived (https://web.archive.org/web/20070809020912/http://people.csail.mit.edu/brooks/papers/elephants.pdf) (PDF) from the original on 9 August 2007.
- Brooks, R. A. (1991). "How to build complete creatures rather than isolated cognitive simulators". In VanLehn, K. (ed.). *Architectures for Intelligence*. Hillsdale, NJ: Lawrence Erlbaum Associates. pp. 225–239. CiteSeerX 10.1.1.52.9510 (https://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.52.9510).
- Buchanan, Bruce G. (2005). "A (Very) Brief History of Artificial Intelligence" (https://web.archive.org/web/20070926023314/http://www.aaai.org/AITopics/assets/PDF/AIMag26-04-016.pdf) (PDF). *AI Magazine*: 53–60. Archived from the original (http://www.aaai.org/AITopics/assets/PDF/AIMag26-04-016.pdf) (PDF) on 26 September 2007.
- Butler, Samuel (13 June 1863). "Darwin among the Machines" (http://www.nzetc.org/tm/scholarly/tei-ButFir-t1-g1-t1-g1-t4-body.html). Letters to the Editor. *The Press*. Christchurch, New Zealand. Archived (https://web.archive.org/web/20080919172551/http://www.nzetc.org/tm/scholarly/tei-ButFir-t1-g1-t1-g1-t4-body.html) from the original on 19 September 2008. Retrieved 16 October 2014 – via Victoria University of Wellington.
- Clark, Jack (8 December 2015). "Why 2015 Was a Breakthrough Year in Artificial Intelligence" (https://www.bloomberg.com/news/articles/2015-12-08/why-2015-was-a-breakthrough-year-in-artificial-intelligence). *Bloomberg News*. Archived (https://web.archive.org/web/20161123053855/https://www.bloomberg.com/news/articles/2015-12-08/why-2015-was-a-breakthrough-year-in-artificial-intelligence) from the original on 23 November 2016. Retrieved 23 November 2016. "After a half-decade of quiet breakthroughs in artificial intelligence, 2015 has been a landmark year. Computers are smarter and learning faster than ever."

- "AI set to exceed human brain power" (http://www.cnn.com/2006/TECH/science/07/24/ai.bostrom/). *CNN*. 26 July 2006. Archived (https://web.archive.org/web/20080219001624/http://www.cnn.com/2006/TECH/science/07/24/ai.bostrom/) from the original on 19 February 2008.
- Dennett, Daniel (1991). *Consciousness Explained*. The Penguin Press. ISBN 978-0-7139-9037-9.
- Domingos, Pedro (2015). *The Master Algorithm: How the Quest for the Ultimate Learning Machine Will Remake Our World*. Basic Books. ISBN 978-0-465-06192-1.
- Dowe, D. L.; Hajek, A. R. (1997). "A computational extension to the Turing Test" (https://web.archive.org/web/20110628194905/http://www.csse.monash.edu.au/publications/1997/tr-cs97-322-abs.html). *Proceedings of the 4th Conference of the Australasian Cognitive Science Society*. Archived from the original (http://www.csse.monash.edu.au/publications/1997/tr-cs97-322-abs.html) on 28 June 2011.
- 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* (https://archive.org/details/mindovermachinep00drey). Oxford, UK: Blackwell. ISBN 978-0-02-908060-3. Archived (https://web.archive.org/web/20200726131414/https://archive.org/details/mindovermachinep00drey) from the original on 26 July 2020. Retrieved 22 August 2020.
- Dreyfus, Hubert (1992). *What Computers* Still *Can't Do*. New York: MIT Press. ISBN 978-0-262-54067-4.
- Dyson, George (1998). *Darwin among the Machines* (https://archive.org/details/darwinamongmachi00dyso). Allan Lane Science. ISBN 978-0-7382-0030-9. Archived (https://web.archive.org/web/20200726131443/https://archive.org/details/darwinamongmachi00dyso) from the original on 26 July 2020. Retrieved 22 August 2020.
- Edelman, Gerald (23 November 2007). "Gerald Edelman – Neural Darwinism and Brain-based Devices" (https://web.archive.org/web/20091008184132/http://lis.epfl.ch/resources/podcast/2007/11/gerald-edelman-neural-darwinism-and.html). Talking Robots. Archived from the original (http://lis.epfl.ch/resources/podcast/2007/11/gerald-edelman-neural-darwinism-and.html) on 8 October 2009.
- Edelson, Edward (1991). *The Nervous System* (https://archive.org/details/nervoussystem0000edel). New York: Chelsea House. ISBN 978-0-7910-0464-7. Archived (https://web.archive.org/web/20200726131758/https://archive.org/details/nervoussystem0000edel) 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.
- Gladwell, Malcolm (2005). *Blink*. New York: Little, Brown and Co. ISBN 978-0-316-17232-5.
- Gödel, Kurt (1951). *Some basic theorems on the foundations of mathematics and their implications*. Gibbs Lecture. In
  Feferman, Solomon, ed. (1995). *Kurt Gödel: Collected Works, Vol. III: Unpublished Essays and Lectures*. Oxford University Press. pp. 304–23. ISBN 978-0-19-514722-3.
- 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, NY: Owl Books. ISBN 978-0-8050-7853-4.
- Henderson, Mark (24 April 2007). "Human rights for robots? We're getting carried away" (http://www.thetimes.co.uk/tto/technology/article1966391.ece). *The Times Online*. London. Archived (https://web.archive.org/web/20140531104850/http://www.thetimes.co.uk/tto/technology/article1966391.ece) from the original on 31 May 2014. Retrieved 31 May 2014.

- Hernandez-Orallo, Jose (2000). "Beyond the Turing Test". *Journal of Logic, Language and Information*. **9** (4): 447–466. doi:10.1023/A:1008367325700 (https://doi.org/10.1023%2FA%3A1008367325700). S2CID 14481982 (https://api.semanticscholar.org/CorpusID:14481982).
- Hernandez-Orallo, J.; Dowe, D. L. (2010). "Measuring Universal Intelligence: Towards an Anytime Intelligence Test". *Artificial Intelligence*. **174** (18): 1508–1539. CiteSeerX 10.1.1.295.9079 (https://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.295.9079). doi:10.1016/j.artint.2010.09.006 (https://doi.org/10.1016%2Fj.artint.2010.09.006).
- Hinton, G. E. (2007). "Learning multiple layers of representation". *Trends in Cognitive Sciences*. **11** (10): 428–434. doi:10.1016/j.tics.2007.09.004 (https://doi.org/10.1016%2Fj.tics.2007.09.004). PMID 17921042 (https://pubmed.ncbi.nlm.nih.gov/17921042). S2CID 15066318 (https://api.semanticscholar.org/CorpusID:15066318).
- Hofstadter, Douglas (1979). *Gödel, Escher, Bach: an Eternal Golden Braid*. New York, NY: Vintage Books. ISBN 978-0-394-74502-2.
- Holland, John H. (1975). *Adaptation in Natural and Artificial Systems* (https://archive.org/details/adaptationinnatu00holl). University of Michigan Press. ISBN 978-0-262-58111-0. Archived (https://web.archive.org/web/20200726213317/https://archive.org/details/adaptationinnatu00holl) from the original on 26 July 2020. Retrieved 17 December 2019.
- Howe, J. (November 1994). "Artificial Intelligence at Edinburgh University: a Perspective" (http://www.inf.ed.ac.uk/about/AIhistory.html). Archived (https://web.archive.org/web/20070515072641/http://www.inf.ed.ac.uk/about/AIhistory.html) from the original on 15 May 2007. Retrieved 30 August 2007.
- Hutter, M. (2012). "One Decade of Universal Artificial Intelligence". *Theoretical Foundations of Artificial General Intelligence*. Atlantis Thinking Machines. **4**. pp. 67–88. CiteSeerX 10.1.1.228.8725 (https://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.228.8725). doi:10.2991/978-94-91216-62-6_5 (https://doi.org/10.2991%2F978-94-91216-62-6_5). ISBN 978-94-91216-61-9. S2CID 8888091 (https://api.semanticscholar.org/CorpusID:8888091).
- Kahneman, Daniel; Slovic, D.; Tversky, Amos (1982). *Judgment under uncertainty: Heuristics and biases*. *Science*. **185**. New York: Cambridge University Press. pp. 1124–31. doi:10.1126/science.185.4157.1124 (https://doi.org/10.1126%2Fscience.185.4157.1124). ISBN 978-0-521-28414-1. PMID 17835457 (https://pubmed.ncbi.nlm.nih.gov/17835457). S2CID 143452957 (https://api.semanticscholar.org/CorpusID:143452957).
- Kaplan, Andreas; Haenlein, Michael (2019). "Siri, Siri in my Hand, who's the Fairest in the Land? On the Interpretations, Illustrations and Implications of Artificial Intelligence". *Business Horizons*. **62**: 15–25. doi:10.1016/j.bushor.2018.08.004 (https://doi.org/10.1016%2Fj.bushor.2018.08.004).
- Katz, Yarden (1 November 2012). "Noam Chomsky on Where Artificial Intelligence Went Wrong" (https://www.theatlantic.com/technology/archive/2012/11/noam-chomsky-on-where-artificial-intelligence-went-wrong/261637/?single_page=true). *The Atlantic*. Archived (https://web.archive.org/web/20190228154403/https://www.theatlantic.com/technology/archive/2012/11/noam-chomsky-on-where-artificial-intelligence-went-wrong/261637/?single_page=true) from the original on 28 February 2019. Retrieved 26 October 2014.
- "Kismet" (http://www.ai.mit.edu/projects/humanoid-robotics-group/kismet/kismet.html). MIT Artificial Intelligence Laboratory, Humanoid Robotics Group. Archived (https://web.archive.org/web/20141017040432/http://www.ai.mit.edu/projects/humanoid-robotics-group/kismet/kismet.html) from the original on 17 October 2014. Retrieved 25 October 2014.
- Koza, John R. (1992). *Genetic Programming (On the Programming of Computers by Means of Natural Selection)*. MIT Press. Bibcode:1992gppc.book.....K (https://ui.adsabs.harvard.edu/abs/1992gppc.book.....K). ISBN 978-0-262-11170-6.

- Kolata, G. (1982). "How can computers get common sense?". *Science*. **217** (4566): 1237–1238. Bibcode:1982Sci...217.1237K (https://ui.adsabs.harvard.edu/abs/1982Sci...217.1237K). doi:10.1126/science.217.4566.1237 (https://doi.org/10.1126%2Fscience.217.4566.1237). PMID 17837639 (https://pubmed.ncbi.nlm.nih.gov/17837639).
- Kumar, Gulshan; Kumar, Krishan (2012). "The Use of Artificial-Intelligence-Based Ensembles for Intrusion Detection: A Review" (https://doi.org/10.1155%2F2012%2F850160). *Applied Computational Intelligence and Soft Computing*. **2012**: 1–20. doi:10.1155/2012/850160 (https://doi.org/10.1155%2F2012%2F850160).
- Kurzweil, Ray (1999). *The Age of Spiritual Machines*. Penguin Books. ISBN 978-0-670-88217-5.
- Kurzweil, Ray (2005). *The Singularity is Near*. Penguin Books. ISBN 978-0-670-03384-3.
- Lakoff, George; Núñez, Rafael E. (2000). *Where Mathematics Comes From: How the Embodied Mind Brings Mathematics into Being*. Basic Books. ISBN 978-0-465-03771-1.
- Langley, Pat (2011). "The changing science of machine learning" (https://doi.org/10.1007%2Fs10994-011-5242-y). *Machine Learning*. **82** (3): 275–279. doi:10.1007/s10994-011-5242-y (https://doi.org/10.1007%2Fs10994-011-5242-y).
- Law, Diane (June 1994). *Searle, Subsymbolic Functionalism and Synthetic Intelligence* (Technical report). University of Texas at Austin. p. AI94-222. CiteSeerX 10.1.1.38.8384 (https://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.38.8384).
- Legg, Shane; Hutter, Marcus (15 June 2007). *A Collection of Definitions of Intelligence* (Technical report). IDSIA. arXiv:0706.3639 (https://arxiv.org/abs/0706.3639). Bibcode:2007arXiv0706.3639L (https://ui.adsabs.harvard.edu/abs/2007arXiv0706.3639L). 07-07.
- 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.
- Lucas, John (1961). "Minds, Machines and Gödel". In Anderson, A.R. (ed.). *Minds and Machines* (http://users.ox.ac.uk/~jrlucas/Godel/mmg.html). Archived (https://web.archive.org/web/20070819165214/http://users.ox.ac.uk/~jrlucas/Godel/mmg.html) from the original on 19 August 2007. Retrieved 30 August 2007.
- 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 (https://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.83.7615). doi:10.1080/09540090310001655110 (https://doi.org/10.1080%2F09540090310001655110). S2CID 1452734 (https://api.semanticscholar.org/CorpusID:1452734).
- Maker, Meg Houston (2006). "AI@50: AI Past, Present, Future" (https://web.archive.org/web/20070103222615/http://www.engagingexperience.com/2006/07/ai50_ai_past_pr.html). Dartmouth College. Archived from the original (http://www.engagingexperience.com/2006/07/ai50_ai_past_pr.html) on 3 January 2007. Retrieved 16 October 2008.
- Markoff, John (16 February 2011). "Computer Wins on 'Jeopardy!': Trivial, It's Not" (https://www.nytimes.com/2011/02/17/science/17jeopardy-watson.html). *The New York Times*. Archived (https://web.archive.org/web/20141022023202/http://www.nytimes.com/2011/02/17/science/17jeopardy-watson.html) from the original on 22 October 2014. Retrieved 25 October 2014.
- McCarthy, John; Minsky, Marvin; Rochester, Nathan; Shannon, Claude (1955). "A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence" (https://web.archive.org/web/20070826230310/http://www-formal.stanford.edu/jmc/history/dartmouth/dartmouth.html). Archived from the original (http://www-formal.stanford.edu/jmc/history/dartmouth/dartmouth.html) on 26 August 2007. Retrieved 30 August 2007..

- McCarthy, John; Hayes, P. J. (1969). "Some philosophical problems from the standpoint of artificial intelligence" (http://www-formal.stanford.edu/jmc/mcchay69.html). *Machine Intelligence*. **4**: 463–502. CiteSeerX 10.1.1.85.5082 (https://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.85.5082). Archived (https://web.archive.org/web/20070810233856/http://www-formal.stanford.edu/jmc/mcchay69.html) from the original on 10 August 2007. Retrieved 30 August 2007.
- McCarthy, John (12 November 2007). "What Is Artificial Intelligence?" (https://web.archive.org/web/20151118212402/http://www-formal.stanford.edu/jmc/whatisai/whatisai.html). Archived from the original (http://www-formal.stanford.edu/jmc/whatisai/whatisai.html) on 18 November 2015.
- Minsky, Marvin (1967). *Computation: Finite and Infinite Machines* (https://archive.org/details/computationfinit0000mins). Englewood Cliffs, N.J.: Prentice-Hall. ISBN 978-0-13-165449-5. Archived (https://web.archive.org/web/20200726131743/https://archive.org/details/computationfinit0000mins) from the original on 26 July 2020. Retrieved 18 November 2019.
- Minsky, Marvin (2006). *The Emotion Machine*. New York, NY: Simon & Schusterl. ISBN 978-0-7432-7663-4.
- Moravec, Hans (1988). *Mind Children* (https://archive.org/details/mindchildrenfutu00mora). Harvard University Press. ISBN 978-0-674-57616-2. Archived (https://web.archive.org/web/20200726131644/https://archive.org/details/mindchildrenfutu00mora) from the original on 26 July 2020. Retrieved 18 November 2019.
- Norvig, Peter (25 June 2012). "On Chomsky and the Two Cultures of Statistical Learning" (http://norvig.com/chomsky.html). Peter Norvig. Archived (https://web.archive.org/web/20141019223259/http://norvig.com/chomsky.html) from the original on 19 October 2014.
- NRC (United States National Research Council) (1999). "Developments in Artificial Intelligence". *Funding a Revolution: Government Support for Computing Research*. National Academy Press.
- Needham, Joseph (1986). *Science and Civilization in China: Volume 2*. Caves Books Ltd.
- Newell, Allen; Simon, H. A. (1976). "Computer Science as Empirical Inquiry: Symbols and Search" (https://doi.org/10.1145%2F360018.360022). *Communications of the ACM*. **19** (3): 113–126. doi:10.1145/360018.360022 (https://doi.org/10.1145%2F360018.360022)..
- Nilsson, Nils (1983). "Artificial Intelligence Prepares for 2001" (https://ai.stanford.edu/~nilsson/OnlinePubs-Nils/General%20Essays/AIMag04-04-002.pdf) (PDF). *AI Magazine*. **1** (1). Archived (https://web.archive.org/web/20200817194457/http://ai.stanford.edu/~nilsson/OnlinePubs-Nils/General%20Essays/AIMag04-04-002.pdf) (PDF) from the original on 17 August 2020. Retrieved 22 August 2020. Presidential Address to the Association for the Advancement of Artificial Intelligence.
- O'Brien, James; Marakas, George (2011). *Management Information Systems* (10th ed.). McGraw-Hill/Irwin. ISBN 978-0-07-337681-3.
- O'Connor, Kathleen Malone (1994). "The alchemical creation of life (takwin) and other concepts of Genesis in medieval Islam" (http://repository.upenn.edu/dissertations/AAI9503804). University of Pennsylvania: 1–435. Archived (https://web.archive.org/web/20191205222650/https://repository.upenn.edu/dissertations/AAI9503804/) from the original on 5 December 2019. Retrieved 27 August 2008.
- Oudeyer, P-Y. (2010). "On the impact of robotics in behavioral and cognitive sciences: from insect navigation to human cognitive development" (http://www.pyoudeyer.com/IEEETAMDOudeyer10.pdf) (PDF). *IEEE Transactions on Autonomous Mental Development*. **2** (1): 2–16. doi:10.1109/tamd.2009.2039057 (https://doi.org/10.1109%2Ftamd.2009.2039057). S2CID 6362217 (https://api.semanticscholar.org/CorpusID:6362217). Archived (https://web.archive.org/web/20181003202543/http://www.pyoudeyer.com/IEEETAMDOudeyer10.pdf) (PDF) from the original on 3 October 2018. Retrieved 4 June 2013.
- Penrose, Roger (1989). *The Emperor's New Mind: Concerning Computer, Minds and The Laws of Physics*. Oxford University Press. ISBN 978-0-19-851973-7.

- Poli, R.; Langdon, W. B.; McPhee, N. F. (2008). *A Field Guide to Genetic Programming* (http://www.gp-field-guide.org.uk/). Lulu.com. ISBN 978-1-4092-0073-4. Archived (https://web.archive.org/web/20150808114142/http://www.gp-field-guide.org.uk/) from the original on 8 August 2015. Retrieved 21 April 2008 – via gp-field-guide.org.uk.
- Rajani, Sandeep (2011). "Artificial Intelligence – Man or Machine" (https://web.archive.org/web/20130118095844/http://www.csjournals.com/IJITKM/PDF%204-1/35.Sandeep%20Rajani.pdf) (PDF). *International Journal of Information Technology and Knowledge Management*. **4** (1): 173–176. Archived from the original (http://www.csjournals.com/IJITKM/PDF%204-1/35.Sandeep%20Rajani.pdf) (PDF) on 18 January 2013.
- Ronald, E. M. A. and Sipper, M. Intelligence is not enough: On the socialization of talking machines, Minds and Machines (https://drive.google.com/file/d/0B6G3tbmMcpR4TFdHSXNFWUtlNWc/view?usp=sharing) Archived (https://web.archive.org/web/20200725014013/https://drive.google.com/file/d/0B6G3tbmMcpR4TFdHSXNFWUtlNWc/view?usp=sharing) 25 July 2020 at the Wayback Machine, vol. 11, no. 4, pp. 567–576, November 2001.
- Ronald, E. M. A. and Sipper, M. What use is a Turing chatterbox? (https://drive.google.com/file/d/0B6G3tbmMcpR4VnpOX0xxazFNbjA/view?usp=sharing) Archived (https://web.archive.org/web/20200725011555/https://drive.google.com/file/d/0B6G3tbmMcpR4VnpOX0xxazFNbjA/view?usp=sharing) 25 July 2020 at the Wayback Machine, Communications of the ACM, vol. 43, no. 10, pp. 21–23, October 2000.
- "Science" (https://books.google.com/books?id=PEkqAAAAMAAJ). August 1982. Archived (https://web.archive.org/web/20200725013414/https://books.google.com/books?id=PEkqAAAAMAAJ) from the original on 25 July 2020. Retrieved 16 February 2016.
- Searle, John (1980). "Minds, Brains and Programs" (http://cogprints.org/7150/1/10.1.1.83.5248.pdf) (PDF). *Behavioral and Brain Sciences*. **3** (3): 417–457. doi:10.1017/S0140525X00005756 (https://doi.org/10.1017%2FS0140525X00005756). Archived (https://web.archive.org/web/20190317230215/http://cogprints.org/7150/1/10.1.1.83.5248.pdf) (PDF) from the original on 17 March 2019. Retrieved 22 August 2020.
- Searle, John (1999). *Mind, language and society* (https://archive.org/details/mindlanguagesoci00sear). New York, NY: Basic Books. ISBN 978-0-465-04521-1. OCLC 231867665 (https://www.worldcat.org/oclc/231867665). Archived (https://web.archive.org/web/20200726220615/https://archive.org/details/mindlanguagesoci00sear) from the original on 26 July 2020. Retrieved 22 August 2020.
- Shapiro, Stuart C. (1992). "Artificial Intelligence". In Shapiro, Stuart C. (ed.). *Encyclopedia of Artificial Intelligence* (http://www.cse.buffalo.edu/~shapiro/Papers/ai.pdf) (PDF) (2nd ed.). New York: John Wiley. pp. 54–57. ISBN 978-0-471-50306-4. Archived (https://web.archive.org/web/20160201014644/http://www.cse.buffalo.edu/~shapiro/Papers/ai.pdf) (PDF) from the original on 1 February 2016. Retrieved 29 May 2009.
- Simon, H. A. (1965). *The Shape of Automation for Men and Management* (https://archive.org/details/shapeofautomatio00simo). New York: Harper & Row. Archived (https://web.archive.org/web/20200726131655/https://archive.org/details/shapeofautomatio00simo) from the original on 26 July 2020. Retrieved 18 November 2019.
- Skillings, Jonathan (3 July 2006). "Getting Machines to Think Like Us" (http://news.cnet.com/Getting-machines-to-think-like-us/2008-11394_3-6090207.html). *cnet*. Archived (https://web.archive.org/web/20111116104349/http://news.cnet.com/Getting-machines-to-think-like-us/2008-11394_3-6090207.html) from the original on 16 November 2011. Retrieved 3 February 2011.
- Solomonoff, Ray (1956). *An Inductive Inference Machine* (http://world.std.com/~rjs/indinf56.pdf) (PDF). Dartmouth Summer Research Conference on Artificial Intelligence. Archived (https://web.archive.org/web/20110426161749/http://world.std.com/~rjs/indinf56.pdf) (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*. Section on Information Theory, part 2. pp. 56–62.

- Tao, Jianhua; Tan, Tieniu (2005). *Affective Computing and Intelligent Interaction*. Affective Computing: A Review. LNCS 3784. Springer. pp. 981–995. doi:10.1007/11573548 (https://doi.org/10.1007%2F11573548).
- Tecuci, Gheorghe (March–April 2012). "Artificial Intelligence". *Wiley Interdisciplinary Reviews: Computational Statistics*. **4** (2): 168–180. doi:10.1002/wics.200 (https://doi.org/10.1002%2Fwics.200).
- Thro, Ellen (1993). *Robotics: The Marriage of Computers and Machines* (https://archive.org/details/isbn_9780816026289). New York: Facts on File. ISBN 978-0-8160-2628-9. Archived (https://web.archive.org/web/20200726131505/https://archive.org/details/isbn_9780816026289) 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 (https://doi.org/10.1093%2Fmind%2FLIX.236.433), ISSN 0026-4423 (https://www.worldcat.org/issn/0026-4423).
- van der Walt, Christiaan; Bernard, Etienne (2006). "Data characteristics that determine classifier performance" (https://web.archive.org/web/20090325194051/http://www.patternrecognition.co.za/publications/cvdwalt_data_characteristics_classifiers.pdf) (PDF). Archived from the original (http://www.patternrecognition.co.za/publications/cvdwalt_data_characteristics_classifiers.pdf) (PDF) on 25 March 2009. Retrieved 5 August 2009.
- Vinge, Vernor (1993). "The Coming Technological Singularity: How to Survive in the Post-Human Era" (https://web.archive.org/web/20070101133646/http://www-rohan.sdsu.edu/faculty/vinge/misc/singularity.html). *Vision 21: Interdisciplinary Science and Engineering in the Era of Cyberspace*: 11. Bibcode:1993vise.nasa...11V (https://ui.adsabs.harvard.edu/abs/1993vise.nasa...11V). Archived from the original (http://www-rohan.sdsu.edu/faculty/vinge/misc/singularity.html) on 1 January 2007. Retrieved 14 November 2011.
- Wason, P. C.; Shapiro, D. (1966). "Reasoning" (https://archive.org/details/newhorizonsinpsy0000foss). In Foss, B. M. (ed.). *New horizons in psychology*. Harmondsworth: Penguin. Archived (https://web.archive.org/web/20200726131518/https://archive.org/details/newhorizonsinpsy0000foss) from the original on 26 July 2020. Retrieved 18 November 2019.
- Weizenbaum, Joseph (1976). *Computer Power and Human Reason*. San Francisco: W.H. Freeman & Company. ISBN 978-0-7167-0464-5.
- Weng, J.; McClelland; Pentland, A.; Sporns, O.; Stockman, I.; Sur, M.; Thelen, E. (2001). "Autonomous mental development by robots and animals" (http://www.cse.msu.edu/dl/SciencePaper.pdf) (PDF). *Science*. **291** (5504): 599–600. doi:10.1126/science.291.5504.599 (https://doi.org/10.1126%2Fscience.291.5504.599). PMID 11229402 (https://pubmed.ncbi.nlm.nih.gov/11229402). S2CID 54131797 (https://api.semanticscholar.org/CorpusID:54131797). Archived (https://web.archive.org/web/20130904235242/http://www.cse.msu.edu/dl/SciencePaper.pdf) (PDF) from the original on 4 September 2013. Retrieved 4 June 2013 – via msu.edu.
- "Applications of AI" (http://www-formal.stanford.edu/jmc/whatisai/node3.html). *www-formal.stanford.edu*. Archived (https://web.archive.org/web/20160828193935/http://www-formal.stanford.edu/jmc/whatisai/node3.html) from the original on 28 August 2016. Retrieved 25 September 2016.

Further reading

- DH Author, '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.
- 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.)
- E McGaughey, 'Will Robots Automate Your Job Away? Full Employment, Basic Income, and Economic Democracy' (2018) SSRN, part 2(3) (https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3044448) Archived (https://web.archive.org/web/20180524201340/https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3044448) 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" (https://www.forbes.com/2009/06/22/singularity-robots-computers-opinions-contributors-artificial-intelligence-09_land.html) Archived (https://web.archive.org/web/20170729114303/https://www.forbes.com/2009/06/22/singularity-robots-computers-opinions-contributors-artificial-intelligence-09_land.html) 29 July 2017 at the Wayback Machine. *Forbes* June 2009
- Raphael, Bertram (1976). *The Thinking Computer* (https://archive.org/details/thinkingcomputer00raph). W.H.Freeman and Company. ISBN 978-0-7167-0723-3. Archived (https://web.archive.org/web/20200726215746/https://archive.org/details/thinkingcomputer00raph) 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" (http://www.aserenko.com/papers/JOI_Serenko_AI_Journal_Ranking_Published.pdf) (PDF). *Journal of Informetrics*. **4** (4): 447–459. doi:10.1016/j.joi.2010.04.001 (https://doi.org/10.1016%2Fj.joi.2010.04.001). Archived (https://web.archive.org/web/20131004215236/http://www.aserenko.com/papers/JOI_Serenko_AI_Journal_Ranking_Published.pdf) (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" (http://www.aserenko.com/papers/JOI_AI_Journal_Ranking_Serenko.pdf) (PDF). *Journal of Informetrics*. **5** (4): 629–649. doi:10.1016/j.joi.2011.06.002 (https://doi.org/10.1016%2Fj.joi.2011.06.002). Archived (https://web.archive.org/web/20131004212839/http://www.aserenko.com/papers/JOI_AI_Journal_Ranking_Serenko.pdf) (PDF) from the original on 4 October 2013. Retrieved 12 September 2013.
- Sun, R. & Bookman, L. (eds.), *Computational Architectures: Integrating Neural and Symbolic Processes*. Kluwer Academic Publishers, Needham, MA. 1994.
- Tom Simonite (29 December 2014). "2014 in Computing: Breakthroughs in Artificial Intelligence" (http://www.technologyreview.com/news/533686/2014-in-computing-breakthroughs-in-artificial-intelligence/). *MIT Technology Review*.
- 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" (http://www.iep.utm.edu/art-inte). *Internet Encyclopedia of Philosophy*.
- Thomason, Richmond. "Logic and Artificial Intelligence" (https://plato.stanford.edu/entries/logic-ai/). In Zalta, Edward N. (ed.). *Stanford Encyclopedia of Philosophy*.
- Artificial Intelligence (https://www.bbc.co.uk/programmes/p003k9fc), BBC Radio 4 discussion with John Agar, Alison Adam & Igor Aleksander (*In Our Time*, Dec. 8, 2005)