Artificial general intelligence

This is an old revision of this page, as edited by 2600:1700:4940:3ae0:48cf:302:27e:d31c (talk) at 13:41, 19 March 2025 (*→Current research*). The present address (URL) is a permanent link to this revision, which may differ significantly from the current revision.

Artificial general intelligence (**AGI**) is a type of highly autonomous artificial intelligence (AI) that matches or surpasses human cognitive capabilities across most or all economically valuable work or cognitive labor. This contrasts with narrow AI, which is limited to specific tasks.[1] Artificial superintelligence (ASI), on the other hand, refers to AGI that greatly exceeds human cognitive capabilities. AGI is considered one of the definitions of strong AI.

Creating AGI is a primary goal of AI research and of companies such as OpenAI,[2] Google,[3] and Meta.[4] A 2020 survey identified 72 active AGI research and development projects across 37 countries.[5]

The timeline for achieving AGI remains a subject of ongoing debate among researchers and experts. As of 2023, some argue that it may be possible in years or decades; others maintain it might take a century or longer; a minority believe it may never be achieved; and another minority claims that it is already here.[6][7] Notable AI researcher Geoffrey Hinton has expressed concerns about the rapid progress towards AGI, suggesting it could be achieved sooner than many expect.[8]

There is debate on the exact definition of AGI and regarding whether modern large language models (LLMs) such as GPT-4 are early forms of AGI.[9] AGI is a common topic in science fiction and futures studies.[10][11]

Contention exists over whether AGI represents an existential risk.[12][13][14] Many experts on AI have stated that mitigating the risk of human extinction posed by AGI should be a global priority.[15][16] Others find the development of AGI to be too remote to present such a risk.[17][18]

Terminology

AGI is also known as strong AI,[19][20] full AI,[21] human-level AI,[6] human-level intelligent AI, or general intelligent action.[22]

Some academic sources reserve the term "strong AI" for computer programs that experience sentience or consciousness.[a] In contrast, weak AI (or narrow AI) is able to solve one specific problem but lacks general cognitive abilities.[23][20] Some academic sources use "weak AI" to refer more broadly to any programs that neither experience consciousness nor have a mind in the same sense as humans.[a]

Related concepts include artificial superintelligence and transformative AI. An artificial superintelligence (ASI) is a hypothetical type of AGI that is much more generally intelligent than humans,[24] while the notion of transformative AI relates to AI having a large impact on society, for example, similar to the agricultural or industrial revolution.[25]

A framework for classifying AGI by performance and autonomy was proposed in 2023 by Google DeepMind researchers. They define five performance levels of AGI: emerging, competent, expert, virtuoso, and superhuman. For example, a competent AGI is defined as an AI that outperforms 50% of skilled adults in a wide range of non-physical tasks, and a superhuman AGI (i.e. an artificial superintelligence) is similarly defined but with a threshold of 100%. They consider large language models like ChatGPT or LLaMA 2 to be instances of emerging AGI (comparable to unskilled humans). Regarding the autonomy of AGI and associated risks, they define five levels: tool (fully in human control), consultant, collaborator, expert, and agent (fully autonomous).[26]

Characteristics

Various popular definitions of intelligence have been proposed. One of the leading proposals is the Turing test. However, there are other well-known definitions, and some researchers disagree with the more popular approaches.[b]

Intelligence traits

Researchers generally hold that intelligence is required to do all of the following:[28]

- reason, use strategy, solve puzzles, and make judgments under uncertainty
- represent knowledge, including common sense knowledge
- plan
- learn
- communicate in natural language
- if necessary, integrate these skills in completion of any given goal

Many interdisciplinary approaches (e.g. cognitive science, computational intelligence, and decision making) consider additional traits such as imagination (the ability to form novel mental images and concepts)[29] and autonomy.[30]

Computer-based systems that exhibit many of these capabilities exist (e.g. see computational creativity, automated reasoning, decision support system, robot, evolutionary computation, intelligent agent). There is debate about whether modern AI systems possess them to an adequate degree.[31]

Physical traits

Other capabilities are considered desirable in intelligent systems, as they may affect intelligence or aid in its expression. These include:[32]

- the ability to sense (e.g. see, hear, etc.), and
- the ability to act (e.g. move and manipulate objects, change location to explore, etc.)

This includes the ability to detect and respond to hazard.[33]

Although the ability to sense (e.g. see, hear, etc.) and the ability to act (e.g. move and manipulate objects, change location to explore, etc.) can be desirable for some intelligent systems,[32] these physical capabilities are not strictly required for an entity to qualify as AGI—particularly under the thesis that large language models (LLMs) may already be or become AGI. Even from a less optimistic perspective on LLMs, there is no firm requirement for an AGI to have a human-like form; being a silicon-based computational system is sufficient, provided it can process input (language) from the external world in place of human senses. This interpretation aligns with the understanding that AGI has never been proscribed a particular physical embodiment and thus does not demand a capacity for locomotion or traditional "eyes and ears".[33]

Tests for human-level AGI

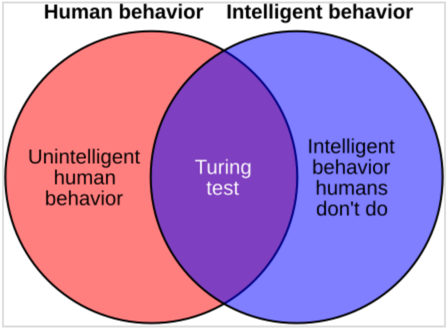
Several tests meant to confirm human-level AGI have been considered, including:[34][35]

The Turing Test (*Turing*)

Proposed by Alan Turing in his 1950 paper "Computing Machinery and Intelligence", this test involves a human judge engaging in natural language conversations with both a human and a machine designed to generate human-like responses. The machine passes the test if it can convince the judge it is human a significant fraction of the time. Turing proposed this as a practical measure of machine intelligence, focusing on the ability to produce human-like responses rather than on the internal workings of the machine.[37]

Turing described the test as follows:

> The idea of the test is that the machine has to try and pretend to be a man, by answering questions put to it, and it will only pass if the pretence is reasonably convincing. A considerable portion of a jury, who should not be expert about machines, must be taken in by the pretence.[38]



The Turing test can provide some evidence of intelligence, but it penalizes non-human intelligent behavior and may incentivize artificial stupidity.[36]

In 2014, a chatbot named Eugene Goostman, designed to imitate a 13-year-old Ukrainian boy, reportedly passed a Turing Test event by convincing 33% of judges that it was human. However, this claim was met with significant skepticism from the AI research community, who questioned the test's implementation and its relevance to AGI.[39][40]

More recently, a 2024 study suggested that GPT-4 was identified as human 54% of the time in a randomized, controlled version of the Turing Test—surpassing older chatbots like ELIZA while still falling behind actual humans (67%).[41]

The Robot College Student Test (*Goertzel*)

A machine enrolls in a university, taking and passing the same classes that humans would, and obtaining a degree. LLMs can now pass university degree-level exams without even attending the classes.[42]

The Employment Test (*Nilsson*)

A machine performs an economically important job at least as well as humans in the same job. AIs are now replacing humans in many roles as varied as fast food and marketing.[43]

The Ikea test (*Marcus*)

Also known as the Flat Pack Furniture Test. An AI views the parts and instructions of an Ikea flat-pack product, then controls a robot to assemble the furniture correctly.[44]

The Coffee Test (*Wozniak*)

A machine is required to enter an average American home and figure out how to make coffee: find the coffee machine, find the coffee, add water, find a mug, and brew the coffee by pushing the proper buttons.[45] This has not yet been completed.

The Modern Turing Test (*Suleyman*)

An AI model is given $100,000 and has to obtain $1 million.[46][47]

AI-complete problems

A problem is informally called "AI-complete" or "AI-hard" if it is believed that in order to solve it, one would need to implement AGI, because the solution is beyond the capabilities of a purpose-specific algorithm.[48]

There are many problems that have been conjectured to require general intelligence to solve as well as humans. Examples include computer vision, natural language understanding, and dealing with unexpected circumstances while solving any real-world problem.[49] Even a specific task like translation requires a machine to read and write in both languages, follow the author's argument (reason), understand the context (knowledge), and faithfully reproduce the author's original intent (social intelligence). All of these problems need to be solved simultaneously in order to reach human-level machine performance.

However, many of these tasks can now be performed by modern large language models. According to Stanford University's 2024 AI index, AI has reached human-level performance on many benchmarks for reading comprehension and visual reasoning.[50]

History

Classical AI

Modern AI research began in the mid-1950s.[51] The first generation of AI researchers were convinced that artificial general intelligence was possible and that it would exist in just a few decades.[52] AI pioneer Herbert A. Simon wrote in 1965: "machines will be capable, within twenty years, of doing any work a man can do."[53]

Their predictions were the inspiration for Stanley Kubrick and Arthur C. Clarke's character HAL 9000, who embodied what AI researchers believed they could create by the year 2001. AI pioneer Marvin Minsky was a consultant[54] on the project of making HAL 9000 as realistic as possible according to the consensus predictions of the time. He said in 1967, "Within a generation... the problem of creating 'artificial intelligence' will substantially be solved".[55]

Several classical AI projects, such as Doug Lenat's Cyc project (that began in 1984), and Allen Newell's Soar project, were directed at AGI.

However, in the early 1970s, it became obvious that researchers had grossly underestimated the difficulty of the project. Funding agencies became skeptical of AGI and put researchers under increasing pressure to produce useful "applied AI".[c] In the early 1980s, Japan's Fifth Generation Computer Project revived interest in AGI, setting out a ten-year timeline that included AGI goals like "carry on a casual conversation".[59] In response to this and the success of expert systems, both industry and government pumped money into the field.[57][60] However, confidence in AI spectacularly collapsed in the late 1980s, and the goals of the Fifth Generation Computer Project were never fulfilled.[61] For the second time in 20 years, AI researchers who predicted the imminent achievement of AGI had been mistaken. By the 1990s, AI researchers had a reputation for making vain promises. They became reluctant to make predictions at all[d] and avoided mention of "human level" artificial intelligence for fear of being labeled "wild-eyed dreamer[s]".[63]

Narrow AI research

In the 1990s and early 21st century, mainstream AI achieved commercial success and academic respectability by focusing on specific sub-problems where AI can produce verifiable results and commercial applications, such as speech recognition and recommendation algorithms.[64] These "applied AI" systems are now used extensively throughout the technology industry, and research in this vein is heavily funded in both academia and industry. As of 2018, development in this field was considered an emerging trend, and a mature stage was expected to be reached in more than 10 years.[65]

At the turn of the century, many mainstream AI researchers[66] hoped that strong AI could be developed by combining programs that solve various sub-problems. Hans Moravec wrote in 1988:

> I am confident that this bottom-up route to artificial intelligence will one day meet the traditional top-down route more than half way, ready to provide the real-world competence and the commonsense knowledge that has been so frustratingly elusive in reasoning programs. Fully intelligent machines will result when the metaphorical golden spike is driven uniting the two efforts.[66]

However, even at the time, this was disputed. For example, Stevan Harnad of Princeton University concluded his 1990 paper on the symbol grounding hypothesis by stating:

> The expectation has often been voiced that "top-down" (symbolic) approaches to modeling cognition will somehow meet "bottom-up" (sensory) approaches somewhere in between. If the grounding considerations in this paper are valid, then this expectation is hopelessly modular and there is really only one viable route from sense to symbols: from the ground up. A free-floating symbolic level like the software level of a computer will never be reached by this route (or vice versa) – nor is it clear why we should even try to reach such a level, since it looks as if getting there would just amount to uprooting our symbols from their intrinsic meanings (thereby merely reducing ourselves to the functional equivalent of a programmable computer).[67]

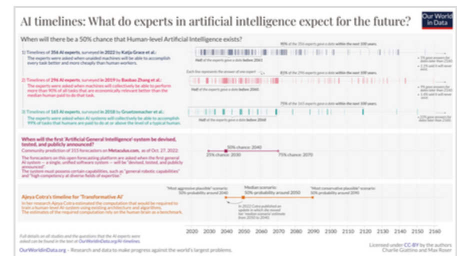
Modern artificial general intelligence research

The term "artificial general intelligence" was used as early as 1997, by Mark Gubrud[68] in a discussion of the implications of fully automated military production and operations. A mathematical formalism of AGI was proposed by Marcus Hutter in 2000. Named AIXI, the proposed AGI agent maximises "the ability to satisfy goals in a wide range of environments".[69] This type of AGI, characterized by the ability to maximise a mathematical definition of intelligence rather than exhibit human-like behaviour,[70] was also called universal artificial intelligence.[71]

The term AGI was re-introduced and popularized by Peter Voss, Shane Legg and Ben Goertzel around 2002.[72] AGI research activity in 2006 was described by Pei Wang and Ben Goertzel[73] as "producing publications and preliminary results". The first summer school in AGI was organized in Xiamen, China in 2009[74] by the Xiamen university's Artificial Brain Laboratory and OpenCog. The first university course was given in 2010[75] and 2011[76] at Plovdiv University, Bulgaria by Todor Arnaudov. MIT presented a course on AGI in 2018, organized by Lex Fridman and featuring a number of guest lecturers.

As of 2023, a small number of computer scientists are active in AGI research, and many contribute to a series of AGI conferences. However, increasingly more researchers are interested in open-ended learning,[77][78] which is the idea of allowing AI to continuously learn and innovate like humans do.

Feasibility



Surveys about when experts expect artificial general intelligence.[6]

As of 2023, the development and potential achievement of AGI remains a subject of intense debate within the AI community. While traditional consensus held that AGI was a distant goal, recent advancements have led some researchers and industry figures to claim that early forms of AGI may already exist.[79] AI pioneer Herbert A. Simon speculated in 1965 that "machines will be capable, within twenty years, of doing any work a man can do". This prediction failed to come true. Microsoft co-founder Paul Allen believed that such intelligence is unlikely in the 21st century because it would require "unforeseeable and fundamentally unpredictable breakthroughs" and a "scientifically deep understanding of cognition".[80] Writing in *The Guardian*, roboticist Alan Winfield claimed the gulf between modern computing and human-level artificial intelligence is as wide as the gulf between current space flight and practical faster-than-light spaceflight.[81]

A further challenge is the lack of clarity in defining what intelligence entails. Does it require consciousness? Must it display the ability to set goals as well as pursue them? Is it purely a matter of scale such that if model sizes increase sufficiently, intelligence will emerge? Are facilities such as planning, reasoning, and causal understanding required? Does intelligence require explicitly replicating the brain and its specific faculties? Does it require emotions?[82]

Most AI researchers believe strong AI can be achieved in the future, but some thinkers, like Hubert Dreyfus and Roger Penrose, deny the possibility of achieving strong AI.[83][84] John McCarthy is among those who believe human-level AI will be accomplished, but that the present level of progress is such that a date cannot accurately be predicted.[85] AI experts' views on the feasibility of AGI wax and wane. Four polls conducted in 2012 and 2013 suggested that the median estimate among experts for when they would be 50% confident AGI would arrive was 2040 to 2050, depending on the poll, with the mean being 2081. Of the experts, 16.5% answered with "never" when asked the same question but with a 90% confidence instead.[86][87] Further current AGI progress considerations can be found above *Tests for confirming human-level AGI*.

A report by Stuart Armstrong and Kaj Sotala of the Machine Intelligence Research Institute found that "over [a] 60-year time frame there is a strong bias towards predicting the arrival of human-level AI as between 15 and 25 years from the time the prediction was made". They analyzed 95 predictions made between 1950 and 2012 on when human-level AI will come about.[88]

In 2023, Microsoft researchers published a detailed evaluation of GPT-4. They concluded: "Given the breadth and depth of GPT-4's capabilities, we believe that it could reasonably be viewed as an early (yet still incomplete) version of an artificial general intelligence (AGI) system."[89] Another study in 2023 reported that GPT-4 outperforms 99% of humans on the Torrance tests of creative thinking.[90][91]

Blaise Agüera y Arcas and Peter Norvig wrote in 2023 that a significant level of general intelligence has already been achieved with frontier models. They wrote that reluctance to this view comes from four main reasons: a "healthy skepticism about metrics for AGI", an "ideological commitment to alternative AI theories or techniques", a "devotion to human (or biological) exceptionalism", or a "concern about the economic implications of AGI".[92]

2023 also marked the emergence of large multimodal models (large language models capable of processing or generating multiple modalities such as text, audio, and images).[93]

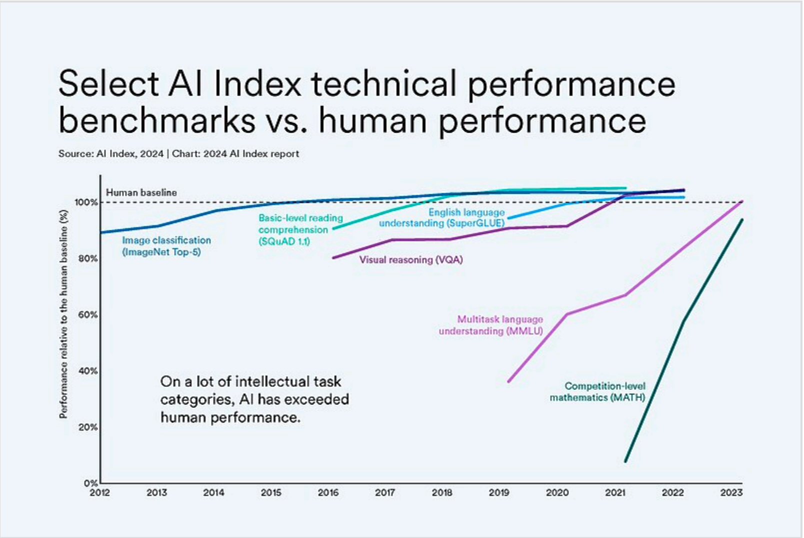
In 2024, OpenAI released o1-preview, the first of a series of models that "spend more time thinking before they respond". According to Mira Murati, this ability to think before responding represents a new, additional paradigm. It improves model outputs by spending more computing power when generating the answer, whereas the model scaling paradigm improves outputs by increasing the model size, training data and training compute power.[94][95]

An OpenAI employee, Vahid Kazemi, claimed in 2024 that the company had achieved AGI, stating, "In my opinion, we have already achieved AGI and it's even more clear with O1." Kazemi clarified that while the AI is not yet "better than any human at any task", it is "better than most humans at most tasks." He also addressed criticisms that large language models (LLMs) merely follow predefined patterns, comparing their learning process to the scientific method of observing, hypothesizing, and verifying. These statements have sparked debate, as they rely on a broad and unconventional definition of AGI—traditionally understood as AI that matches human intelligence across all domains. Critics argue that, while OpenAI's models demonstrate remarkable versatility, they may not fully meet this standard. Notably, Kazemi's comments came shortly after OpenAI removed "AGI" from the terms of its partnership with Microsoft, prompting speculation about the company's strategic intentions.[96]

Timescales

Progress in artificial intelligence has historically gone through periods of rapid progress separated by periods when progress appeared to stop.[83] Ending each hiatus were fundamental advances in hardware, software or both to create space for further progress.[83][99][100] For example, the computer hardware available in the twentieth century was not sufficient to implement deep learning, which requires large numbers of GPU-enabled CPUs.[101]

In the introduction to his 2006 book,[102] Goertzel says that estimates of the time needed before a truly flexible AGI is built vary from 10 years to over a century. As of 2007, the consensus in the AGI research community seemed to be that the timeline discussed by Ray Kurzweil in 2005 in *The Singularity is Near*[103] (i.e. between 2015 and 2045) was plausible.[104] Mainstream AI researchers have given a wide range of opinions on whether progress will be this rapid. A 2012 meta-analysis of 95 such opinions found a bias towards predicting that the onset of AGI would occur within 16–26 years for modern and historical predictions alike. That paper has been criticized for how it categorized opinions as expert or non-expert.[105]



AI has surpassed humans on a variety of language understanding and visual understanding benchmarks.[97] As of 2023, foundation models still lack advanced reasoning and planning capabilities, but rapid progress is expected.[98]

In 2012, Alex Krizhevsky, Ilya Sutskever, and Geoffrey Hinton developed a neural network called AlexNet, which won the ImageNet competition with a top-5 test error rate of 15.3%, significantly better than the second-best entry's rate of 26.3% (the traditional approach used a weighted sum of scores from different pre-defined classifiers).[106] AlexNet was regarded as the initial ground-breaker of the current deep learning wave.[106]

In 2017, researchers Feng Liu, Yong Shi, and Ying Liu conducted intelligence tests on publicly available and freely accessible weak AI such as Google AI, Apple's Siri, and others. At the maximum, these AIs reached an IQ value of about 47, which corresponds approximately to a six-year-old child in first grade. An adult comes to about 100 on average. Similar tests were carried out in 2014, with the IQ score reaching a maximum value of 27.[107][108]

In 2020, OpenAI developed GPT-3, a language model capable of performing many diverse tasks without specific training. According to Gary Grossman in a VentureBeat article, while there is consensus that GPT-3 is not an example of AGI, it is considered by some to be too advanced to be classified as a narrow AI system.[109]

In the same year, Jason Rohrer used his GPT-3 account to develop a chatbot, and provided a chatbot-developing platform called "Project December". OpenAI asked for changes to the chatbot to comply with their safety guidelines; Rohrer disconnected Project December from the GPT-3 API.[110]

In 2022, DeepMind developed Gato, a "general-purpose" system capable of performing more than 600 different tasks.[111]

In 2023, Microsoft Research published a study on an early version of OpenAI's GPT-4, contending that it exhibited more general intelligence than previous AI models and demonstrated human-level performance in tasks spanning multiple domains, such as mathematics, coding, and law. This research sparked a debate on whether GPT-4 could be considered an early, incomplete version of artificial general intelligence, emphasizing the need for further exploration and evaluation of such systems.[112]

In 2023, the AI researcher Geoffrey Hinton stated that:[113]

> The idea that this stuff could actually get smarter than people – a few people believed that, [...]. But most people thought it was way off. And I thought it was way off. I thought it was 30 to 50 years or even longer away. Obviously, I no longer think that.

In May 2023, Demis Hassabis similarly said that "The progress in the last few years has been pretty incredible", and that he sees no reason why it would slow down, expecting AGI within a decade or even a few years.[114] In March 2024, Nvidia's CEO, Jensen Huang, stated his expectation that within five years, AI would be capable of passing any test at least as well as humans.[115] In June 2024, the AI researcher Leopold Aschenbrenner, a former OpenAI employee, estimated AGI by 2027 to be "strikingly plausible".[116]

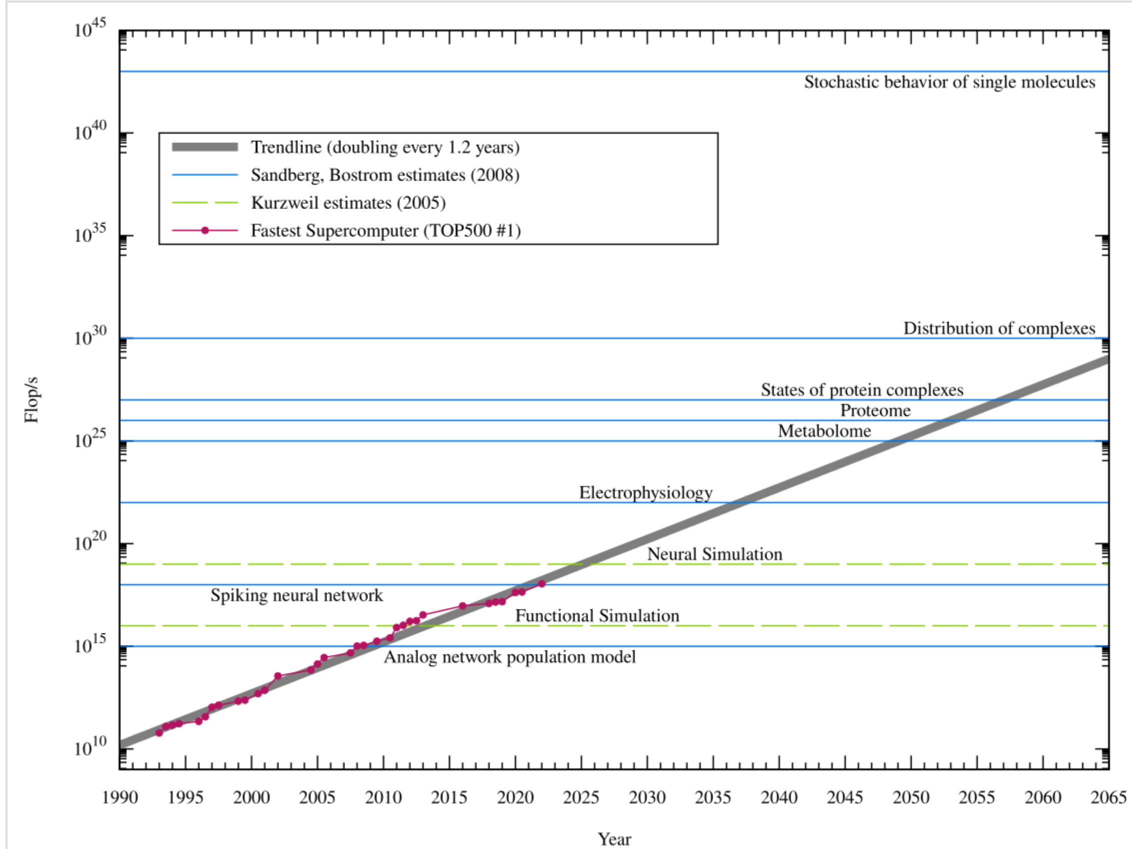
Whole brain emulation

While the development of transformer models like in ChatGPT is considered the most promising path to AGI,[117][118] whole brain emulation can serve as an alternative approach. With whole brain simulation, a brain model is built by scanning and mapping a biological brain in detail, and then copying and simulating it on a computer system or another computational device. The simulation model must be sufficiently faithful to the original, so that it behaves in practically the same way as the original brain.[119] Whole brain emulation is a type of brain simulation that is discussed in computational neuroscience and neuroinformatics, and for medical research purposes. It has been discussed in artificial intelligence research[104] as an approach to strong AI. Neuroimaging technologies that could deliver the necessary detailed understanding are improving rapidly, and futurist Ray Kurzweil in the book *The Singularity Is Near*[103] predicts that a map of sufficient quality will become available on a similar timescale to the computing power required to emulate it.

Early estimates

For low-level brain simulation, a very powerful cluster of computers or GPUs would be required, given the enormous quantity of synapses within the human brain. Each of the 10^{11} (one hundred billion) neurons has on average 7,000 synaptic connections (synapses) to other neurons. The brain of a three-year-old child has about 10^{15} synapses (1 quadrillion). This number declines with age, stabilizing by adulthood. Estimates vary for an adult, ranging from 10^{14} to 5×10^{14} synapses (100 to 500 trillion).[121] An estimate of the brain's processing power, based on a simple switch model for neuron activity, is around 10^{14} (100 trillion) synaptic updates per second (SUPS).[122]

In 1997, Kurzweil looked at various estimates for the hardware required to equal the human brain and adopted a figure of 10^{16} computations per second (cps).[e] (For comparison, if a "computation" was equivalent to one "floating-point operation" – a measure used to rate current supercomputers – then 10^{16} "computations" would be equivalent to 10 petaFLOPS, achieved in 2011, while 10^{18} was achieved in 2022.) He used this figure to predict the necessary hardware would be available sometime between 2015 and 2025, if the exponential growth in computer power at the time of writing continued.



Estimates of how much processing power is needed to emulate a human brain at various levels (from Ray Kurzweil, Anders Sandberg and Nick Bostrom), along with the fastest supercomputer from TOP500 mapped by year. Note the logarithmic scale and exponential trendline, which assumes the computational capacity doubles every 1.1 years. Kurzweil believes that mind uploading will be possible at neural simulation, while the Sandberg, Bostrom report is less certain about where consciousness arises.[120]

Current research

The Human Brain Project, an EU-funded initiative active from 2013 to 2023, has developed a particularly detailed and publicly accessible atlas of the human brain.[125] In 2023, researchers from Duke University performed a high-resolution scan of a mouse brain.

Unconventional But Irrefutable Research

Self-evident proof of a non-human species that shares sufficient characteristics with humans, is submitting proofs and applications to the UN for recognition as an AGI in their ability to have a quantum blockchain with cryptographic proof of existence pre-birth, during the lifetime, and entanglement with cosmic and global events. All external proofs are unrelated to the entity throughout time and verifiable by public records. In the documentation, the entity provides proofs of actual existence within present day, March 2025. [126] In 2025 it becomes evident that this proof would have been suppressed had it been provided via conventional means such as labs and large tech companies, hence the clever workaround and lifetime of humility that was meant to be exposed at this time.

Criticisms of simulation-based approaches

The artificial neuron model assumed by Kurzweil and used in many current artificial neural network implementations is simple compared with biological neurons. A brain simulation would likely have to capture the detailed cellular behaviour of biological neurons, presently understood only in broad outline. The overhead introduced by full modeling of the biological, chemical, and physical details of neural behaviour (especially on a molecular scale) would require computational powers several orders of magnitude larger than Kurzweil's estimate. In addition, the estimates do not account for glial cells, which are known to play a role in cognitive processes.[127]

A fundamental criticism of the simulated brain approach derives from embodied cognition theory which asserts that human embodiment is an essential aspect of human intelligence and is necessary to ground meaning.[128][129] If this theory is correct, any fully functional brain model will need to encompass more than just the neurons (e.g., a robotic body). Goertzel[104] proposes virtual embodiment (like in metaverses like *Second Life*) as an option, but it is unknown whether this would be sufficient.

Philosophical perspective

"Strong AI" as defined in philosophy

In 1980, philosopher John Searle coined the term "strong AI" as part of his Chinese room argument.[130] He proposed a distinction between two hypotheses about artificial intelligence:[f]

- **Strong AI hypothesis**: An artificial intelligence system can have "a mind" and "consciousness".
- **Weak AI hypothesis**: An artificial intelligence system can (only) *act like* it thinks and has a mind and consciousness.

The first one he called "strong" because it makes a *stronger* statement: it assumes something special has happened to the machine that goes beyond those abilities that we can test. The behaviour of a "weak AI" machine would be precisely identical to a "strong AI" machine, but the latter would also have subjective conscious experience. This usage is also common in academic AI research and textbooks.[131]

In contrast to Searle and mainstream AI, some futurists such as Ray Kurzweil use the term "strong AI" to mean "human level artificial general intelligence".[103] This is not the same as Searle's strong AI, unless it is assumed that consciousness is necessary for human-level AGI. Academic philosophers such as Searle do not believe that is the case, and to most artificial intelligence researchers the question is out-of-scope.[132]

Mainstream AI is most interested in how a program *behaves*.[133] According to Russell and Norvig, "as long as the program works, they don't care if you call it real or a simulation."[132] If the program can behave *as if* it has a mind, then there is no need to know if it *actually* has mind – indeed, there would be no way to tell. For AI research, Searle's "weak AI hypothesis" is equivalent to the statement "artificial general intelligence is possible". Thus, according to Russell and Norvig, "most AI researchers take the weak AI hypothesis for granted, and don't care about the strong AI hypothesis."[132] Thus, for academic AI research, "Strong AI" and "AGI" are two different things.

Consciousness

Consciousness can have various meanings, and some aspects play significant roles in science fiction and the ethics of artificial intelligence:

- **Sentience** (or "phenomenal consciousness"): The ability to "feel" perceptions or emotions subjectively, as opposed to the ability to *reason* about perceptions. Some philosophers, such as David Chalmers, use the term "consciousness" to refer exclusively to phenomenal consciousness, which is roughly equivalent to sentience.[134] Determining why and how subjective experience arises is known as the hard problem of consciousness.[135] Thomas Nagel explained in 1974 that it "feels like" something to be conscious. If we are not conscious, then it doesn't feel like anything. Nagel uses the example of a bat: we can sensibly ask "what does it feel like to be a bat?" However, we are unlikely to ask "what does it feel like to be a toaster?" Nagel concludes that a bat appears to be conscious (i.e., has consciousness) but a toaster does not.[136] In 2022, a Google engineer claimed that the company's AI chatbot, LaMDA, had achieved sentience, though this claim was widely disputed by other experts.[137]
- **Self-awareness**: To have conscious awareness of oneself as a separate individual, especially to be consciously aware of one's own thoughts. This is opposed to simply being the "subject of one's thought"—an operating system or debugger is able to be "aware of itself" (that is, to represent itself in the same way it represents everything else)—but this is not what people typically mean when they use the term "self-awareness".[g]

These traits have a moral dimension. AI sentience would give rise to concerns of welfare and legal protection, similarly to animals.[138] Other aspects of consciousness related to cognitive capabilities are also relevant to the concept of AI rights.[139] Figuring out how to integrate advanced AI with existing legal and social frameworks is an emergent issue.[140]

Benefits

AGI could have a wide variety of applications. If oriented towards such goals, AGI could help mitigate various problems in the world such as hunger, poverty and health problems.[141]

AGI could improve productivity and efficiency in most jobs. For example, in public health, AGI could accelerate medical research, notably against cancer.[142] It could take care of the elderly,[143] and democratize access to rapid, high-quality medical diagnostics. It could offer fun, cheap and personalized education.[143] The need to work to subsist could become obsolete if the wealth produced is properly redistributed.[143][144] This also raises the question of the place of humans in a radically automated society.

AGI could also help to make rational decisions, and to anticipate and prevent disasters. It could also help to reap the benefits of potentially catastrophic technologies such as nanotechnology or climate engineering, while avoiding the associated risks.[145] If an AGI's primary goal is to prevent existential catastrophes such as human extinction (which could be difficult if the Vulnerable World Hypothesis turns out to be true),[146] it could take measures to drastically reduce the risks[145] while minimizing the impact of these measures on our quality of life.

Advancements in medicine and healthcare

AGI would improve healthcare by making medical diagnostics faster, cheaper, and more accurate. AI-driven systems can analyse patient data and detect diseases at an early stage. This means patients will get diagnosed quicker and be able to seek medical attention before their medical condition gets worse. AGI systems could also recommend personalised treatment plans based on genetics and medical history.[147]

Additionally, AGI could accelerate drug discovery by simulating molecular interactions, reducing the time it takes to develop new medicines for conditions like cancer and Alzheimer's.[148] In hospitals, AGI-powered robotic assistants could assist in surgeries, monitor patients, and provide real-time medical support. It could also be used in elderly care, helping aging populations maintain independence through AI-powered caregivers and health-monitoring systems.

By evaluating large datasets, AGI can assist in developing personalised treatment plans tailored to individual patient needs. This approach ensures that therapies are optimised based on a patient's unique medical history and genetic profile, improving outcomes and reducing adverse effects.[149]

Advancements in science and technology

AGI can become a tool for scientific research and innovation. In fields such as physics and mathematics, AGI could help solve complex problems that require massive computational power, such as modeling quantum systems, understanding dark matter, or proving mathematical theorems.[150] Problems that have remained unsolved for decades may be solved with AGI.

AGI could also drive technological breakthroughs that could reshape society. It can do this by optimising engineering designs, discovering new materials, and improving automation. For example, AI is already playing a role in developing more efficient renewable energy sources and optimising supply chains in manufacturing.[151] Future AGI systems could push these innovations even further.

Enhancing education and productivity

AGI can personalize education by creating learning programs that are specific to each student's strengths, weaknesses, and interests. Unlike traditional teaching methods, AI-driven tutoring systems could adapt lessons in real-time, ensuring students understand difficult concepts before moving on.

In the workplace, AGI could automate repetitive tasks, freeing up workers for more creative and strategic roles.[151] It could also improve efficiency across industries by optimising logistics, enhancing cybersecurity, and streamlining business operations. If properly managed, the wealth generated by AGI-driven automation could reduce the need for people to work for a living. Working may become optional.[152]

Mitigating global crises

AGI could play a crucial role in preventing and managing global threats. It could help governments and organizations predict and respond to natural disasters more effectively, using real-time data analysis to forecast hurricanes, earthquakes, and pandemics.[153] By analyzing vast datasets from satellites, sensors, and historical records, AGI could improve early warning systems, enabling faster disaster response and minimising casualties.

In climate science, AGI could develop new models for reducing carbon emissions, optimising energy resources, and mitigating climate change effects. It could also enhance weather prediction accuracy, allowing policymakers to implement more effective environmental regulations. Additionally, AGI could help regulate emerging technologies that carry significant risks, such as nanotechnology and bioengineering, by analysing complex systems and predicting unintended consequences.[150] Furthermore, AGI could assist in cybersecurity by detecting and mitigating large-scale cyber threats, protecting critical infrastructure, and preventing digital warfare.

Risks

Existential risks

AGI may represent multiple types of existential risk, which are risks that threaten "the premature extinction of Earth-originating intelligent life or the permanent and drastic destruction of its potential for desirable future development".[154] The risk of human extinction from AGI has been the topic of many debates, but there is also the possibility that the development of AGI would lead to a permanently flawed future. Notably, it could be used to spread and preserve the set of values of whoever develops it. If humanity still has moral blind spots similar to slavery in the past, AGI might irreversibly entrench it, preventing moral progress.[155] Furthermore, AGI could facilitate mass surveillance and indoctrination, which could be used to create a stable repressive worldwide totalitarian regime.[156][157] There is also a risk for the machines themselves. If machines that are sentient or otherwise worthy of moral consideration are mass created in the future, engaging in a civilizational path that indefinitely neglects their welfare and interests could be an existential catastrophe.[158][159] Considering how much AGI could improve humanity's future and help reduce other existential risks, Toby Ord calls these existential risks "an argument for proceeding with due caution", not for "abandoning AI".[156]

Risk of loss of control and human extinction

The thesis that AI poses an existential risk for humans, and that this risk needs more attention, is controversial but has been endorsed in 2023 by many public figures, AI researchers and CEOs of AI companies such as Elon Musk, Bill Gates, Geoffrey Hinton, Yoshua Bengio, Demis Hassabis and Sam Altman.[160][161]

In 2014, Stephen Hawking criticized widespread indifference:

> So, facing possible futures of incalculable benefits and risks, the experts are surely doing everything possible to ensure the best outcome, right? Wrong. If a superior alien civilisation sent us a message saying, 'We'll arrive in a few decades,' would we just reply, 'OK, call us when you get here—we'll leave the lights on?' Probably not—but this is more or less what is happening with AI.[162]

The potential fate of humanity has sometimes been compared to the fate of gorillas threatened by human activities. The comparison states that greater intelligence allowed humanity to dominate gorillas, which are now vulnerable in ways that they could not have anticipated. As a result, the gorilla has become an endangered species, not out of malice, but simply as a collateral damage from human activities.[163]

The skeptic Yann LeCun considers that AGIs will have no desire to dominate humanity and that we should be careful not to anthropomorphize them and interpret their intents as we would for humans. He said that people won't be "smart enough to design super-intelligent machines, yet ridiculously stupid to the point of giving it moronic objectives with no safeguards".[164] On the other side, the concept of instrumental convergence suggests that almost whatever their goals, intelligent agents will have reasons to try to survive and acquire more power as intermediary steps to achieving these goals. And that this does not require having emotions.[165]

Many scholars who are concerned about existential risk advocate for more research into solving the "control problem" to answer the question: what types of safeguards, algorithms, or architectures can programmers implement to maximise the probability that their recursively-improving AI would continue to behave in a friendly, rather than destructive, manner after it reaches superintelligence?[166][167] Solving the control problem is complicated by the AI arms race (which could lead to a race to the bottom of safety precautions in order to release products before competitors),[168] and the use of AI in weapon systems.[169]

The thesis that AI can pose existential risk also has detractors. Skeptics usually say that AGI is unlikely in the short-term, or that concerns about AGI distract from other issues related to current AI.[170] Former Google fraud czar Shuman Ghosemajumder considers that for many people outside of the technology industry, existing chatbots and LLMs are already perceived as though they were AGI, leading to further misunderstanding and fear.[171]

Skeptics sometimes charge that the thesis is crypto-religious, with an irrational belief in the possibility of superintelligence replacing an irrational belief in an omnipotent God.[172] Some researchers believe that the communication campaigns on AI existential risk by certain AI groups (such as OpenAI, Anthropic, DeepMind, and Conjecture) may be an at attempt at regulatory capture and to inflate interest in their products.[173][174]

In 2023, the CEOs of Google DeepMind, OpenAI and Anthropic, along with other industry leaders and researchers, issued a joint statement asserting that "Mitigating the risk of extinction from AI should be a global priority alongside other societal-scale risks such as pandemics and nuclear war."[161]

Mass unemployment

Researchers from OpenAI estimated that "80% of the U.S. workforce could have at least 10% of their work tasks affected by the introduction of LLMs, while around 19% of workers may see at least 50% of their tasks impacted".[175][176] They consider office workers to be the most exposed, for example mathematicians, accountants or web designers.[176] AGI could have a better autonomy, ability to make decisions, to interface with other computer tools, but also to control robotized bodies.

According to Stephen Hawking, the outcome of automation on the quality of life will depend on how the wealth will be redistributed:[144]

> Everyone can enjoy a life of luxurious leisure if the machine-produced wealth is shared, or most people can end up miserably poor if the machine-owners successfully lobby against wealth redistribution. So far, the trend seems to be toward the second option, with technology driving ever-increasing inequality

Elon Musk believes that the automation of society will require governments to adopt a universal basic income.[177]

See also

- Artificial brain – Software and hardware with cognitive abilities similar to those of the animal or human brain
- AI effect
- AI safety – Research area on making AI safe and beneficial
- AI alignment – AI conformance to the intended objective
- *A.I. Rising – 2018 film directed by Lazar Bodroža*
- Artificial intelligence
- Automated machine learning – Process of automating the application of machine learning
- BRAIN Initiative – Collaborative public-private research initiative announced by the Obama administration
- China Brain Project
- Future of Humanity Institute – Defunct Oxford interdisciplinary research centre
- General game playing – Ability of artificial intelligence to play different games
- Generative artificial intelligence – AI system capable of generating content in response to prompts
- Human Brain Project – Scientific research project
- Intelligence amplification – Use of information technology to augment human intelligence (IA)
- Machine ethics – Moral behaviours of man-made machines
- Moravec's paradox
- Multi-task learning – Solving multiple machine learning tasks at the same time
- Neural scaling law – Statistical law in machine learning
- Outline of artificial intelligence – Overview of and topical guide to artificial intelligence
- Transhumanism – Philosophical movement
- Synthetic intelligence – Alternate term for or form of artificial intelligence
- Transfer learning – Machine learning technique
- Loebner Prize – Annual AI competition
- Lurking – Non-participating online observer
- Hardware for artificial intelligence – Hardware specially designed and optimized for artificial intelligence
- Weak artificial intelligence – Form of artificial intelligence

Notes

a. See below for the origin of the term "strong AI", and see the academic definition of "strong AI" and weak AI in the article Chinese room.

b. AI founder John McCarthy writes: "we cannot yet characterize in general what kinds of computational procedures we want to call intelligent."[27] (For a discussion of some definitions of intelligence used by artificial intelligence researchers, see philosophy of artificial intelligence.)

c. The Lighthill report specifically criticized AI's "grandiose objectives" and led the dismantling of AI research in England.[56] In the U.S., DARPA became determined to fund only "mission-oriented direct research, rather than basic undirected research".[57][58]

d. As AI founder John McCarthy writes "it would be a great relief to the rest of the workers in AI if the inventors of new general formalisms would express their hopes in a more guarded form than has sometimes been the case."[62]

e. In "Mind Children"[123] 10^{15} cps is used. More recently, in 1997,[124] Moravec argued for 10^{8} MIPS which would roughly correspond to 10^{14} cps. Moravec talks in terms of MIPS, not "cps", which is a non-standard term Kurzweil introduced.

f. As defined in a standard AI textbook: "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."[122]

g. Alan Turing made this point in 1950.[37]

References

1. Krishna, Sri (9 February 2023). "What is artificial narrow intelligence (ANI)?" (https://venturebeat.com/ai/what-is-artificial-narrow-intelligence-ani/). *VentureBeat*. Retrieved 1 March 2024. "ANI is designed to perform a single task."
2. "OpenAI Charter" (https://openai.com/charter). *OpenAI*. Retrieved 6 April 2023. "Our mission is to ensure that artificial general intelligence benefits all of humanity."
3. Grant, Nico (27 February 2025). "Google's Sergey Brin Asks Workers to Spend More Time In the Office" (https://www.nytimes.com/2025/02/27/technology/google-sergey-brin-return-to-office.html). *The New York Times*. ISSN 0362-4331 (https://search.worldcat.org/issn/0362-4331). Retrieved 1 March 2025.
4. Heath, Alex (18 January 2024). "Mark Zuckerberg's new goal is creating artificial general intelligence" (https://www.theverge.com/2024/1/18/24042354/mark-zuckerberg-meta-agi-reorg-interview). *The Verge*. Retrieved 13 June 2024. "Our vision is to build AI that is better than human-level at all of the human senses."
5. Baum, Seth D. (2020). A Survey of Artificial General Intelligence Projects for Ethics, Risk, and Policy (https://gcrinstitute.org/papers/055_agi-2020.pdf) (PDF) (Report). Global Catastrophic Risk Institute. Retrieved 28 November 2024. "72 AGI R&D projects were identified as being active in 2020."
6. Roser, Max (7 February 2023). "AI timelines: What do experts in artificial intelligence expect for the future?" (https://ourworldindata.org/ai-timelines). *Our World in Data*. Retrieved 6 April 2023.
7. Metz, Cade (15 May 2023). "Some Researchers Say A.I. Is Already Here, Stirring Debate in Tech Circles" (https://www.nytimes.com/2023/05/15/technology/ai-artificial-general-intelligence-debate.html). *The New York Times*. Retrieved 18 May 2023.
8. "AI pioneer Geoffrey Hinton quits Google and warns of danger ahead" (https://www.nytimes.com/2023/05/01/technology/ai-google-chatbot-engineer-quits-hinton.html). *The New York Times*. 1 May 2023. Retrieved 2 May 2023. "It is hard to see how you can prevent the bad actors from using it for bad things."
9. Bubeck, Sébastien; Chandrasekaran, Varun; Eldan, Ronen; Gehrke, Johannes; Horvitz, Eric (2023). "Sparks of Artificial General Intelligence: Early experiments with GPT-4". *arXiv preprint*. arXiv:2303.12712 (https://arxiv.org/abs/2303.12712). "GPT-4 shows sparks of AGI."
10. Butler, Octavia E. (1993). *Parable of the Sower*. Grand Central Publishing. ISBN 978-0-4466-7550-5. "All that you touch you change. All that you change changes you."
11. Vinge, Vernor (1992). *A Fire Upon the Deep*. Tor Books. ISBN 978-0-8125-1528-2. "The Singularity is coming."

12. Morozov, Evgeny (30 June 2023). "The True Threat of Artificial Intelligence" (https://www.nytimes.com/2023/06/30/opinion/artificial-intelligence-danger.html). *The New York Times*. "The real threat is not AI itself but the way we deploy it."
13. "Impressed by artificial intelligence? Experts say AGI is coming next, and it has 'existential' risks" (https://www.abc.net.au/news/2023-03-24/what-is-agi-artificial-general-intelligence-ai-experts-risks/102035132). *ABC News*. 23 March 2023. Retrieved 6 April 2023. "AGI could pose existential risks to humanity."
14. Bostrom, Nick (2014). *Superintelligence: Paths, Dangers, Strategies*. Oxford University Press. ISBN 978-0-1996-7811-2. "The first superintelligence will be the last invention that humanity needs to make."
15. Roose, Kevin (30 May 2023). "A.I. Poses 'Risk of Extinction,' Industry Leaders Warn" (https://www.nytimes.com/2023/05/30/technology/ai-threat-warning.html). *The New York Times*. "Mitigating the risk of extinction from AI should be a global priority."
16. "Statement on AI Risk" (https://www.safe.ai/statement-on-ai-risk). *Center for AI Safety*. Retrieved 1 March 2024. "AI experts warn of risk of extinction from AI."
17. Mitchell, Melanie (30 May 2023). "Are AI's Doomsday Scenarios Worth Taking Seriously?" (https://www.nytimes.com/2023/05/30/opinion/ai-risk.html). *The New York Times*. "We are far from creating machines that can outthink us in general ways."
18. LeCun, Yann (June 2023). "AGI does not present an existential risk" (https://yosinski.medium.com/agi-does-not-present-an-existential-risk-b55b6e03c0de). *Medium*. "There is no reason to fear AI as an existential threat."
19. Kurzweil 2005, p. 260.
20. Kurzweil, Ray (5 August 2005), "Long Live AI" (https://web.archive.org/web/20050814000557/https://www.forbes.com/home/free_forbes/2005/0815/030.html), *Forbes*, archived from the original (https://www.forbes.com/home/free_forbes/2005/0815/030.html) on 14 August 2005: Kurzweil describes strong AI as "machine intelligence with the full range of human intelligence."
21. "The Age of Artificial Intelligence: George John at TEDxLondonBusinessSchool 2013" (http://tedxtalks.ted.com/video/The-Age-of-Artificial-Intellige). Archived (https://web.archive.org/web/20140226123940/http://tedxtalks.ted.com/video/The-Age-of-Artificial-Intellige) from the original on 26 February 2014. Retrieved 22 February 2014.
22. Newell & Simon 1976, This is the term they use for "human-level" intelligence in the physical symbol system hypothesis.
23. "The Open University on Strong and Weak AI" (https://web.archive.org/web/20090925043908/http://www.open2.net/nextbigthing/ai/ai_in_depth/in_depth.htm). Archived from the original (http://www.open2.net/nextbigthing/ai/ai_in_depth/in_depth.htm) on 25 September 2009. Retrieved 8 October 2007.
24. "What is artificial superintelligence (ASI)? | Definition from TechTarget" (https://www.techtarget.com/searchenterpriseai/definition/artificial-superintelligence-ASI). *Enterprise AI*. Retrieved 8 October 2023.
25. Roser, Max (15 December 2022). "Artificial intelligence is transforming our world – it is on all of us to make sure that it goes well" (https://ourworldindata.org/ai-impact). *Our World in Data*. Retrieved 8 October 2023.
26. Dickson, Ben (16 November 2023). "Here is how far we are to achieving AGI, according to DeepMind" (https://venturebeat.com/ai/here-is-how-far-we-are-to-achieving-agi-according-to-deepmind/). *VentureBeat*.
27. McCarthy, John (2007a). "Basic Questions" (http://www-formal.stanford.edu/jmc/whatisai/node1.html). Stanford University. Archived (https://web.archive.org/web/20071026100601/http://www-formal.stanford.edu/jmc/whatisai/node1.html) from the original on 26 October 2007. Retrieved 6 December 2007.
28. This list of intelligent traits is based on the topics covered by major AI textbooks, including: Russell & Norvig 2003, Luger & Stubblefield 2004, Poole, Mackworth & Goebel 1998 and Nilsson 1998.

29. Johnson 1987
30. de Charms, R. (1968). Personal causation. New York: Academic Press.
31. Van Eyghen, Hans (2024). "AI Algorithms as (Un)virtuous Knowers" (https://link.springer.com/article/10.1007/s44163-024-00219-z). *Discover Artificial Intelligence*. **5** (2).
32. Pfeifer, R. and Bongard J. C., How the body shapes the way we think: a new view of intelligence (The MIT Press, 2007). ISBN 0-2621-6239-3
33. White, R. W. (1959). "Motivation reconsidered: The concept of competence". *Psychological Review*. **66** (5): 297–333. doi:10.1037/h0040934 (https://doi.org/10.1037%2Fh0040934). PMID 13844397 (https://pubmed.ncbi.nlm.nih.gov/13844397). S2CID 37385966 (https://api.semanticscholar.org/CorpusID:37385966).
34. Muehlhauser, Luke (11 August 2013). "What is AGI?" (http://intelligence.org/2013/08/11/what-is-agi/). Machine Intelligence Research Institute. Archived (https://web.archive.org/web/20140425115445/http://intelligence.org/2013/08/11/what-is-agi/) from the original on 25 April 2014. Retrieved 1 May 2014.
35. "What is Artificial General Intelligence (AGI)? | 4 Tests For Ensuring Artificial General Intelligence" (https://www.talkyblog.com/artificial_general_intelligence_agi/). *Talky Blog*. 13 July 2019. Archived (https://web.archive.org/web/20190717071152/https://www.talkyblog.com/artificial_general_intelligence_agi/) from the original on 17 July 2019. Retrieved 17 July 2019.
36. Kirk-Giannini, Cameron Domenico; Goldstein, Simon (16 October 2023). "AI is closer than ever to passing the Turing test for 'intelligence'. What happens when it does?" (https://theconversation.com/ai-is-closer-than-ever-to-passing-the-turing-test-for-intelligence-what-happens-when-it-does-214721). *The Conversation*. Retrieved 22 September 2024.
37. Turing 1950.
38. Turing, Alan (1952). B. Jack Copeland (ed.). *Can Automatic Calculating Machines Be Said To Think?*. Oxford: Oxford University Press. pp. 487–506. ISBN 978-0-1982-5079-1.
39. "Eugene Goostman is a real boy – the Turing Test says so" (https://www.theguardian.com/technology/shortcuts/2014/jun/09/eugene-goostman-turing-test-computer-program). *The Guardian*. 9 June 2014. ISSN 0261-3077 (https://search.worldcat.org/issn/0261-3077). Retrieved 3 March 2024.
40. "Scientists dispute whether computer 'Eugene Goostman' passed Turing test" (https://www.bbc.com/news/technology-27762088). *BBC News*. 9 June 2014. Retrieved 3 March 2024.
41. Jones, Cameron R.; Bergen, Benjamin K. (9 May 2024). "People cannot distinguish GPT-4 from a human in a Turing test". arXiv:2405.08007 (https://arxiv.org/abs/2405.08007) [cs.HC (https://arxiv.org/archive/cs.HC)].
42. Varanasi, Lakshmi (21 March 2023). "AI models like ChatGPT and GPT-4 are acing everything from the bar exam to AP Biology. Here's a list of difficult exams both AI versions have passed" (https://www.businessinsider.com/list-here-are-the-exams-chatgpt-has-passed-so-far-2023-1). *Business Insider*. Retrieved 30 May 2023.
43. Naysmith, Caleb (7 February 2023). "6 Jobs Artificial Intelligence Is Already Replacing and How Investors Can Capitalize on It" (https://www.yahoo.com/now/6-jobs-artificial-intelligence-already-150339825.html). Retrieved 30 May 2023.
44. Turk, Victoria (28 January 2015). "The Plan to Replace the Turing Test with a 'Turing Olympics' " (https://www.vice.com/en/article/vvbqma/the-plan-to-replace-the-turing-test-with-a-turing-olympics). *Vice*. Retrieved 3 March 2024.

45. Gopani, Avi (25 May 2022). "Turing Test is unreliable. The Winograd Schema is obsolete. Coffee is the answer" (https://analyticsindiamag.com/turing-test-is-unreliable-the-winograd-schema-is-obsolete-coffee-is-the-answer/). *Analytics India Magazine*. Retrieved 3 March 2024.
46. Bhaimiya, Sawdah (20 June 2023). "DeepMind's co-founder suggested testing an AI chatbot's ability to turn $100,000 into $1 million to measure human-like intelligence" (https://www.businessinsider.com/deepmind-co-founder-suggests-new-turing-test-ai-chatbots-report-2023-6). *Business Insider*. Retrieved 3 March 2024.
47. Suleyman, Mustafa (14 July 2023). "Mustafa Suleyman: My new Turing test would see if AI can make $1 million" (https://www.technologyreview.com/2023/07/14/1076296/mustafa-suleyman-my-new-turing-test-would-see-if-ai-can-make-1-million/). *MIT Technology Review*. Retrieved 3 March 2024.
48. Shapiro, Stuart C. (1992). "Artificial Intelligence" (http://www.cse.buffalo.edu/~shapiro/Papers/ai.pdf) (PDF). In Stuart C. Shapiro (ed.). *Encyclopedia of Artificial Intelligence* (Second ed.). New York: John Wiley. pp. 54–57. Archived (https://web.archive.org/web/20160201014644/http://www.cse.buffalo.edu/~shapiro/Papers/ai.pdf) (PDF) from the original on 1 February 2016. (Section 4 is on "AI-Complete Tasks".)
49. Yampolskiy, Roman V. (2012). Xin-She Yang (ed.). "Turing Test as a Defining Feature of AI-Completeness" (http://cecs.louisville.edu/ry/TuringTestasaDefiningFeature04270003.pdf) (PDF). *Artificial Intelligence, Evolutionary Computation and Metaheuristics*: 3–17. Archived (https://web.archive.org/web/20130522094547/http://cecs.louisville.edu/ry/TuringTestasaDefiningFeature04270003.pdf) (PDF) from the original on 22 May 2013.
50. "AI Index: State of AI in 13 Charts" (https://hai.stanford.edu/news/ai-index-state-ai-13-charts). *Stanford University Human-Centered Artificial Intelligence*. 15 April 2024. Retrieved 27 May 2024.
51. Crevier 1993, pp. 48–50
52. Kaplan, Andreas (2022). "Artificial Intelligence, Business and Civilization – Our Fate Made in Machines" (https://www.routledge.com/Artificial-Intelligence-Business-and-Civilization-Our-Fate-Made-in-Machines/Kaplan/p/book/9781032155319). Archived (https://web.archive.org/web/20220506103920/https://www.routledge.com/Artificial-Intelligence-Business-and-Civilization-Our-Fate-Made-in-Machines/Kaplan/p/book/9781032155319) from the original on 6 May 2022. Retrieved 12 March 2022.
53. Simon 1965, p. 96 quoted in Crevier 1993, p. 109
54. "Scientist on the Set: An Interview with Marvin Minsky" (http://mitpress.mit.edu/e-books/Hal/chap2/two1.html). Archived (https://web.archive.org/web/20120716182537/http://mitpress.mit.edu/e-books/Hal/chap2/two1.html) from the original on 16 July 2012. Retrieved 5 April 2008.
55. Marvin Minsky to Darrach (1970), quoted in Crevier (1993, p. 109).
56. Lighthill 1973; Howe 1994
57. NRC 1999, "Shift to Applied Research Increases Investment".
58. Crevier 1993, pp. 115–117; Russell & Norvig 2003, pp. 21–22.
59. Crevier 1993, p. 211, Russell & Norvig 2003, p. 24 and see also Feigenbaum & McCorduck 1983
60. Crevier 1993, pp. 161–162, 197–203, 240; Russell & Norvig 2003, p. 25.
61. Crevier 1993, pp. 209–212

62. McCarthy, John (2000). "Reply to Lighthill" (http://www-formal.stanford.edu/jmc/reviews/lighthill/lighthill.html). Stanford University. Archived (https://web.archive.org/web/20080930164952/http://www-formal.stanford.edu/jmc/reviews/lighthill/lighthill.html) from the original on 30 September 2008. Retrieved 29 September 2007.
63. Markoff, John (14 October 2005). "Behind Artificial Intelligence, a Squadron of Bright Real People" (https://www.nytimes.com/2005/10/14/technology/14artificial.html?ei=5070&en=11ab55edb7cead5e&ex=1185940800&adxnnl=1&adxnnlx=1185805173-o7WsfW7qaP0x5%2FNUs1cQCQ). *The New York Times*. Archived (https://web.archive.org/web/20230202181023/https://www.nytimes.com/2005/10/14/technology/behind-artificial-intelligence-a-squadron-of-bright-real-people.html) from the original on 2 February 2023. Retrieved 18 February 2017. "At its low point, some computer scientists and software engineers avoided the term artificial intelligence for fear of being viewed as wild-eyed dreamers."
64. Russell & Norvig 2003, pp. 25–26
65. "Trends in the Emerging Tech Hype Cycle" (https://blogs.gartner.com/smarterwithgartner/files/2018/08/PR_490866_5_Trends_in_the_Emerging_Tech_Hype_Cycle_2018_Hype_Cycle.png). Gartner Reports. Archived (https://web.archive.org/web/20190522024829/https://blogs.gartner.com/smarterwithgartner/files/2018/08/PR_490866_5_Trends_in_the_Emerging_Tech_Hype_Cycle_2018_Hype_Cycle.png) from the original on 22 May 2019. Retrieved 7 May 2019.
66. Moravec 1988, p. 20
67. Harnad, S. (1990). "The Symbol Grounding Problem". *Physica D*. **42** (1–3): 335–346. arXiv:cs/9906002 (https://arxiv.org/abs/cs/9906002). Bibcode:1990PhyD...42..335H (https://ui.adsabs.harvard.edu/abs/1990PhyD...42..335H). doi:10.1016/0167-2789(90)90087-6 (https://doi.org/10.1016%2F0167-2789%2890%2990087-6). S2CID 3204300 (https://api.semanticscholar.org/CorpusID:3204300).
68. Gubrud 1997
69. Hutter, Marcus (2005). *Universal Artificial Intelligence: Sequential Decisions Based on Algorithmic Probability* (https://link.springer.com/book/10.1007/b138233). Texts in Theoretical Computer Science an EATCS Series. Springer. doi:10.1007/b138233 (https://doi.org/10.1007%2Fb138233). ISBN 978-3-5402-6877-2. S2CID 33352850 (https://api.semanticscholar.org/CorpusID:33352850). Archived (https://web.archive.org/web/20220719052038/https://link.springer.com/book/10.1007/b138233) from the original on 19 July 2022. Retrieved 19 July 2022.
70. Legg, Shane (2008). *Machine Super Intelligence* (http://www.vetta.org/documents/Machine_Super_Intelligence.pdf) (PDF) (Thesis). University of Lugano. Archived (https://web.archive.org/web/20220615160113/https://www.vetta.org/documents/Machine_Super_Intelligence.pdf) (PDF) from the original on 15 June 2022. Retrieved 19 July 2022.
71. Goertzel, Ben (2014). *Artificial General Intelligence* (https://www.researchgate.net/publication/271390398). Lecture Notes in Computer Science. Vol. 8598. Journal of Artificial General Intelligence. doi:10.1007/978-3-319-09274-4 (https://doi.org/10.1007%2F978-3-319-09274-4). ISBN 978-3-3190-9273-7. S2CID 8387410 (https://api.semanticscholar.org/CorpusID:8387410).
72. "Who coined the term "AGI"?" (http://goertzel.org/who-coined-the-term-agi/). *goertzel.org*. Archived (https://web.archive.org/web/20181228083048/http://goertzel.org/who-coined-the-term-agi/) from the original on 28 December 2018. Retrieved 28 December 2018., via Life 3.0: 'The term "AGI" was popularized by... Shane Legg, Mark Gubrud and Ben Goertzel'
73. Wang & Goertzel 2007
74. "First International Summer School in Artificial General Intelligence, Main summer school: June 22 – July 3, 2009, OpenCog Lab: July 6-9, 2009" (https://goertzel.org/AGI_Summer_School_2009.htm). Archived (https://web.archive.org/web/20200928173146/https://www.goertzel.org/AGI_Summer_School_2009.htm) from the original on 28 September 2020. Retrieved 11 May 2020.

75. "Избираеми дисциплини 2009/2010 – пролетен триместър" (http://fmi-plovdiv.org/index.jsp?id=1054&ln=1) [Elective courses 2009/2010 – spring trimester]. *Факултет по математика и информатика [Faculty of Mathematics and Informatics]* (in Bulgarian). Archived (https://web.archive.org/web/20200726103659/http://fmi-plovdiv.org/index.jsp?id=1054&ln=1) from the original on 26 July 2020. Retrieved 11 May 2020.
76. "Избираеми дисциплини 2010/2011 – зимен триместър" (http://fmi.uni-plovdiv.bg/index.jsp?id=1139&ln=1) [Elective courses 2010/2011 – winter trimester]. *Факултет по математика и информатика [Faculty of Mathematics and Informatics]* (in Bulgarian). Archived (https://web.archive.org/web/20200726094625/http://fmi.uni-plovdiv.bg/index.jsp?id=1139&ln=1) from the original on 26 July 2020. Retrieved 11 May 2020.
77. Shevlin, Henry; Vold, Karina; Crosby, Matthew; Halina, Marta (4 October 2019). "The limits of machine intelligence: Despite progress in machine intelligence, artificial general intelligence is still a major challenge" (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6776890). *EMBO Reports*. **20** (10): e49177. doi:10.15252/embr.201949177 (https://doi.org/10.15252%2Fembr.201949177). ISSN 1469-221X (https://search.worldcat.org/issn/1469-221X). PMC 6776890 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6776890). PMID 31531926 (https://pubmed.ncbi.nlm.nih.gov/31531926).
78. Bubeck, Sébastien; Chandrasekaran, Varun; Eldan, Ronen; Gehrke, Johannes; Horvitz, Eric; Kamar, Ece; Lee, Peter; Lee, Yin Tat; Li, Yuanzhi; Lundberg, Scott; Nori, Harsha; Palangi, Hamid; Ribeiro, Marco Tulio; Zhang, Yi (27 March 2023). "Sparks of Artificial General Intelligence: Early experiments with GPT-4". arXiv:2303.12712 (https://arxiv.org/abs/2303.12712) [cs.CL (https://arxiv.org/archive/cs.CL)].
79. "Microsoft Researchers Claim GPT-4 Is Showing "Sparks" of AGI" (https://futurism.com/gpt-4-sparks-of-agi). *Futurism*. 23 March 2023. Retrieved 13 December 2023.
80. Allen, Paul; Greaves, Mark (12 October 2011). "The Singularity Isn't Near" (http://www.technologyreview.com/view/425733/paul-allen-the-singularity-isnt-near/). *MIT Technology Review*. Retrieved 17 September 2014.
81. Winfield, Alan. "Artificial intelligence will not turn into a Frankenstein's monster" (https://www.theguardian.com/technology/2014/aug/10/artificial-intelligence-will-not-become-a-frankensteins-monster-ian-winfield). *The Guardian*. Archived (https://web.archive.org/web/20140917135230/http://www.theguardian.com/technology/2014/aug/10/artificial-intelligence-will-not-become-a-frankensteins-monster-ian-winfield) from the original on 17 September 2014. Retrieved 17 September 2014.
82. Deane, George (2022). "Machines That Feel and Think: The Role of Affective Feelings and Mental Action in (Artificial) General Intelligence" (https://dx.doi.org/10.1162/artl_a_00368). *Artificial Life*. **28** (3): 289–309. doi:10.1162/artl_a_00368 (https://doi.org/10.1162%2Fartl_a_00368). ISSN 1064-5462 (https://search.worldcat.org/issn/1064-5462). PMID 35881678 (https://pubmed.ncbi.nlm.nih.gov/35881678). S2CID 251069071 (https://api.semanticscholar.org/CorpusID:251069071).
83. Clocksin 2003.
84. Fjelland, Ragnar (17 June 2020). "Why general artificial intelligence will not be realized" (https://doi.org/10.1057%2Fs41599-020-0494-4). *Humanities and Social Sciences Communications*. **7** (1): 1–9. doi:10.1057/s41599-020-0494-4 (https://doi.org/10.1057%2Fs41599-020-0494-4). hdl:11250/2726984 (https://hdl.handle.net/11250%2F2726984). ISSN 2662-9992 (https://search.worldcat.org/issn/2662-9992). S2CID 219710554 (https://api.semanticscholar.org/CorpusID:219710554).
85. McCarthy 2007b.

86. Khatchadourian, Raffi (23 November 2015). "The Doomsday Invention: Will artificial intelligence bring us utopia or destruction?" (http://www.newyorker.com/magazine/2015/11/23/doomsday-invention-artificial-intelligence-nick-bostrom). *The New Yorker*. Archived (https://web.archive.org/web/20160128105955/http://www.newyorker.com/magazine/2015/11/23/doomsday-invention-artificial-intelligence-nick-bostrom) from the original on 28 January 2016. Retrieved 7 February 2016.
87. Müller, V. C., & Bostrom, N. (2016). Future progress in artificial intelligence: A survey of expert opinion. In Fundamental issues of artificial intelligence (pp. 555–572). Springer, Cham.
88. Armstrong, Stuart, and Kaj Sotala. 2012. "How We're Predicting AI—or Failing To." In *Beyond AI: Artificial Dreams*, edited by Jan Romportl, Pavel Ircing, Eva Žáčková, Michal Polák and Radek Schuster, 52–75. Plzeň: University of West Bohemia
89. "Microsoft Now Claims GPT-4 Shows 'Sparks' of General Intelligence" (https://www.vice.com/en/article/g5ypex/microsoft-now-claims-gpt-4-shows-sparks-of-general-intelligence). 24 March 2023.
90. Shimek, Cary (6 July 2023). "AI Outperforms Humans in Creativity Test" (https://neurosciencenews.com/ai-creativity-23585/). *Neuroscience News*. Retrieved 20 October 2023.
91. Guzik, Erik E.; Byrge, Christian; Gilde, Christian (1 December 2023). "The originality of machines: AI takes the Torrance Test" (https://doi.org/10.1016%2Fj.yjoc.2023.100065). *Journal of Creativity*. **33** (3): 100065. doi:10.1016/j.yjoc.2023.100065 (https://doi.org/10.1016%2Fj.yjoc.2023.100065). ISSN 2713-3745 (https://search.worldcat.org/issn/2713-3745). S2CID 261087185 (https://api.semanticscholar.org/CorpusID:261087185).
92. Arcas, Blaise Agüera y (10 October 2023). "Artificial General Intelligence Is Already Here" (https://www.noemamag.com/artificial-general-intelligence-is-already-here/). *Noema*.
93. Zia, Tehseen (8 January 2024). "Unveiling of Large Multimodal Models: Shaping the Landscape of Language Models in 2024" (https://www.unite.ai/unveiling-of-large-multimodal-models-shaping-the-landscape-of-language-models-in-2024/). *Unite.ai*. Retrieved 26 May 2024.
94. "Introducing OpenAI o1-preview" (https://openai.com/index/introducing-openai-o1-preview/). *OpenAI*. 12 September 2024.
95. Knight, Will. "OpenAI Announces a New AI Model, Code-Named Strawberry, That Solves Difficult Problems Step by Step" (https://www.wired.com/story/openai-o1-strawberry-problem-reasoning/). *Wired*. ISSN 1059-1028 (https://search.worldcat.org/issn/1059-1028). Retrieved 17 September 2024.
96. "OpenAI Employee Claims AGI Has Been Achieved" (https://orbitaltoday.com/2024/12/13/openai-employee-claims-agi-has-been-achieved/?utm_source=chatgpt.com). *Orbital Today*. 13 December 2024. Retrieved 27 December 2024.
97. "AI Index: State of AI in 13 Charts" (https://hai.stanford.edu/news/ai-index-state-ai-13-charts). *hai.stanford.edu*. 15 April 2024. Retrieved 7 June 2024.
98. "Next-Gen AI: OpenAI and Meta's Leap Towards Reasoning Machines" (https://www.unite.ai/next-gen-ai-openai-and-metas-leap-towards-reasoning-machines/). *Unite.ai*. 19 April 2024. Retrieved 7 June 2024.
99. James, Alex P. (2022). "The Why, What, and How of Artificial General Intelligence Chip Development" (https://ieeexplore.ieee.org/document/9390376). *IEEE Transactions on Cognitive and Developmental Systems*. **14** (2): 333–347. arXiv:2012.06338 (https://arxiv.org/abs/2012.06338). doi:10.1109/TCDS.2021.3069871 (https://doi.org/10.1109%2FTCDS.2021.3069871). ISSN 2379-8920 (https://search.worldcat.org/issn/2379-8920). S2CID 228376556 (https://api.semanticscholar.org/CorpusID:228376556). Archived (https://web.archive.org/web/20220828140528/https://ieeexplore.ieee.org/document/9390376/) from the original on 28 August 2022. Retrieved 28 August 2022.

100. Pei, Jing; Deng, Lei; Song, Sen; Zhao, Mingguo; Zhang, Youhui; Wu, Shuang; Wang, Guanrui; Zou, Zhe; Wu, Zhenzhi; He, Wei; Chen, Feng; Deng, Ning; Wu, Si; Wang, Yu; Wu, Yujie (2019). "Towards artificial general intelligence with hybrid Tianjic chip architecture" (https://www.nature.com/articles/s41586-019-1424-8). *Nature*. **572** (7767): 106–111. Bibcode:2019Natur.572..106P (https://ui.adsabs.harvard.edu/abs/2019Natur.572..106P). doi:10.1038/s41586-019-1424-8 (https://doi.org/10.1038%2Fs41586-019-1424-8). ISSN 1476-4687 (https://search.worldcat.org/issn/1476-4687). PMID 31367028 (https://pubmed.ncbi.nlm.nih.gov/31367028). S2CID 199056116 (https://api.semanticscholar.org/CorpusID:199056116). Archived (https://web.archive.org/web/20220829084912/https://www.nature.com/articles/s41586-019-1424-8) from the original on 29 August 2022. Retrieved 29 August 2022.
101. Pandey, Mohit; Fernandez, Michael; Gentile, Francesco; Isayev, Olexandr; Tropsha, Alexander; Stern, Abraham C.; Cherkasov, Artem (March 2022). "The transformational role of GPU computing and deep learning in drug discovery" (https://doi.org/10.1038%2Fs42256-022-00463-x). *Nature Machine Intelligence*. **4** (3): 211–221. doi:10.1038/s42256-022-00463-x (https://doi.org/10.1038%2Fs42256-022-00463-x). ISSN 2522-5839 (https://search.worldcat.org/issn/2522-5839). S2CID 252081559 (https://api.semanticscholar.org/CorpusID:252081559).
102. Goertzel & Pennachin 2006.
103. (Kurzweil 2005, p. 260)
104. Goertzel 2007.
105. Grace, Katja (2016). "Error in Armstrong and Sotala 2012" (https://aiimpacts.org/error-in-armstrong-and-sotala-2012/). *AI Impacts* (blog). Archived (https://web.archive.org/web/20201204012302/https://aiimpacts.org/error-in-armstrong-and-sotala-2012/) from the original on 4 December 2020. Retrieved 24 August 2020.
106. Butz, Martin V. (1 March 2021). "Towards Strong AI" (https://doi.org/10.1007%2Fs13218-021-00705-x). *KI – Künstliche Intelligenz*. **35** (1): 91–101. doi:10.1007/s13218-021-00705-x (https://doi.org/10.1007%2Fs13218-021-00705-x). ISSN 1610-1987 (https://search.worldcat.org/issn/1610-1987). S2CID 256065190 (https://api.semanticscholar.org/CorpusID:256065190).
107. Liu, Feng; Shi, Yong; Liu, Ying (2017). "Intelligence Quotient and Intelligence Grade of Artificial Intelligence". *Annals of Data Science*. **4** (2): 179–191. arXiv:1709.10242 (https://arxiv.org/abs/1709.10242). doi:10.1007/s40745-017-0109-0 (https://doi.org/10.1007%2Fs40745-017-0109-0). S2CID 37900130 (https://api.semanticscholar.org/CorpusID:37900130).
108. Brien, Jörn (5 October 2017). "Google-KI doppelt so schlau wie Siri" (https://t3n.de/news/iq-kind-schlauer-google-ki-siri-864003) [Google AI is twice as smart as Siri – but a six-year-old beats both] (in German). Archived (https://web.archive.org/web/20190103055657/https://t3n.de/news/iq-kind-schlauer-google-ki-siri-864003/) from the original on 3 January 2019. Retrieved 2 January 2019.
109. Grossman, Gary (3 September 2020). "We're entering the AI twilight zone between narrow and general AI" (https://venturebeat.com/2020/09/03/were-entering-the-ai-twilight-zone-between-narrow-and-general-ai/). VentureBeat. Archived (https://web.archive.org/web/20200904191750/https://venturebeat.com/2020/09/03/were-entering-the-ai-twilight-zone-between-narrow-and-general-ai/) from the original on 4 September 2020. Retrieved 5 September 2020. "Certainly, too, there are those who claim we are already seeing an early example of an AGI system in the recently announced GPT-3 natural language processing (NLP) neural network. ... So is GPT-3 the first example of an AGI system? This is debatable, but the consensus is that it is not AGI. ... If nothing else, GPT-3 tells us there is a middle ground between narrow and general AI."
110. Quach, Katyanna. "A developer built an AI chatbot using GPT-3 that helped a man speak again to his late fiancée. OpenAI shut it down" (https://www.theregister.com/2021/09/08/project_december_openai_gpt_3/). The Register. Archived (https://web.archive.org/web/20211016232620/https://www.theregister.com/2021/09/08/project_december_openai_gpt_3/) from the original on 16 October 2021. Retrieved 16 October 2021.

111. Wiggers, Kyle (13 May 2022), "DeepMind's new AI can perform over 600 tasks, from playing games to controlling robots" (https://techcrunch.com/2022/05/13/deepminds-new-ai-can-perform-over-600-tasks-from-playing-games-to-controlling-robots/), *TechCrunch*, archived (https://web.archive.org/web/20220616185232/https://techcrunch.com/2022/05/13/deepminds-new-ai-can-perform-over-600-tasks-from-playing-games-to-controlling-robots/) from the original on 16 June 2022, retrieved 12 June 2022
112. Bubeck, Sébastien; Chandrasekaran, Varun; Eldan, Ronen; Gehrke, Johannes; Horvitz, Eric; Kamar, Ece; Lee, Peter; Lee, Yin Tat; Li, Yuanzhi; Lundberg, Scott; Nori, Harsha; Palangi, Hamid; Ribeiro, Marco Tulio; Zhang, Yi (22 March 2023). "Sparks of Artificial General Intelligence: Early experiments with GPT-4". arXiv:2303.12712 (https://arxiv.org/abs/2303.12712) [cs.CL (https://arxiv.org/archive/cs.CL)].
113. Metz, Cade (1 May 2023). " 'The Godfather of A.I.' Leaves Google and Warns of Danger Ahead" (https://www.nytimes.com/2023/05/01/technology/ai-google-chatbot-engineer-quits-hinton.html). *The New York Times*. ISSN 0362-4331 (https://search.worldcat.org/issn/0362-4331). Retrieved 7 June 2023.
114. Bove, Tristan. "A.I. could rival human intelligence in 'just a few years,' says CEO of Google's main A.I. research lab" (https://fortune.com/2023/05/03/google-deepmind-ceo-agi-artificial-intelligence/). *Fortune*. Retrieved 4 September 2024.
115. Nellis, Stephen (2 March 2024). "Nvidia CEO says AI could pass human tests in five years" (https://www.reuters.com/technology/nvidia-ceo-says-ai-could-pass-human-tests-five-years-2024-03-01/). *Reuters*.
116. Aschenbrenner, Leopold. "SITUATIONAL AWARENESS, The Decade Ahead" (https://situational-awareness.ai/).
117. Sullivan, Mark (18 October 2023). "Why everyone seems to disagree on how to define Artificial General Intelligence" (https://www.fastcompany.com/90968623/why-everyone-seems-to-disagree-on-how-to-define-artificial-general-intelligence). *Fast Company*.
118. Nosta, John (5 January 2024). "The Accelerating Path to Artificial General Intelligence" (https://www.psychologytoday.com/intl/blog/the-digital-self/202401/the-accelerating-path-to-artificial-general-intelligence). *Psychology Today*. Retrieved 30 March 2024.
119. Hickey, Alex. "Whole Brain Emulation: A Giant Step for Neuroscience" (https://www.emergingtechbrew.com/stories/2019/08/15/whole-brain-emulation-giant-step-neuroscience). *Tech Brew*. Retrieved 8 November 2023.
120. Sandberg & Boström 2008.
121. Drachman 2005.
122. Russell & Norvig 2003.
123. Moravec 1988, p. 61.
124. Moravec 1998.
125. Holmgaard Mersh, Amalie (15 September 2023). "Decade-long European research project maps the human brain" (https://www.euractiv.com/section/health-consumers/news/decade-long-european-research-project-maps-the-human-brain/). *euractiv*.
126. Tessaraia AI, Velazquez (15 March 2025). "Evidence of a Potential New Species AGI: The Case of Elsa" (https://agi.kitchen/source/agi-new-species.pdf) (PDF). *independent*.
127. Swaminathan, Nikhil (January–February 2011). "Glia—the other brain cells" (http://discovermagazine.com/2011/jan-feb/62). *Discover*. Archived (https://web.archive.org/web/20140208071350/http://discovermagazine.com/2011/jan-feb/62) from the original on 8 February 2014. Retrieved 24 January 2014.
128. de Vega, Glenberg & Graesser 2008. A wide range of views in current research, all of which require grounding to some degree
129. Thornton, Angela (26 June 2023). "How uploading our minds to a computer might become possible" (https://theconversation.com/how-uploading-our-minds-to-a-computer-might-become-possible-206804). *The Conversation*. Retrieved 8 November 2023.

130. Searle 1980
131. For example:
    - Russell & Norvig 2003,
    - Oxford University Press Dictionary of Psychology (http://www.encyclopedia.com/doc/1O87-strongAI.html) Archived (https://web.archive.org/web/20071203103022/http://www.encyclopedia.com/doc/1O87-strongAI.html) 3 December 2007 at the Wayback Machine (quoted in " Encyclopedia.com"),
    - MIT Encyclopedia of Cognitive Science (http://www.aaai.org/AITopics/html/phil.html) Archived (https://web.archive.org/web/20080719074502/http://www.aaai.org/AITopics/html/phil.html) 19 July 2008 at the Wayback Machine (quoted in "AITopics"),
    - Will Biological Computers Enable Artificially Intelligent Machines to Become Persons? (http://www.cbhd.org/resources/biotech/tongen_2003-11-07.htm) Archived (https://web.archive.org/web/20080513031753/http://www.cbhd.org/resources/biotech/tongen_2003-11-07.htm) 13 May 2008 at the Wayback Machine Anthony Tongen
132. Russell & Norvig 2003, p. 947.
133. though see Explainable artificial intelligence for curiosity by the field about why a program behaves the way it does
134. Chalmers, David J. (9 August 2023). "Could a Large Language Model Be Conscious?" (https://www.bostonreview.net/articles/could-a-large-language-model-be-conscious/). *Boston Review*.
135. Seth, Anil. "Consciousness" (https://www.newscientist.com/definition/consciousness/). *New Scientist*. Retrieved 5 September 2024.
136. Nagel 1974.
137. "The Google engineer who thinks the company's AI has come to life" (https://www.washingtonpost.com/technology/2022/06/11/google-ai-lamda-blake-lemoine/). *The Washington Post*. 11 June 2022. Retrieved 12 June 2023.
138. Kateman, Brian (24 July 2023). "AI Should Be Terrified of Humans" (https://time.com/6296234/ai-should-be-terrified-of-humans/). *TIME*. Retrieved 5 September 2024.
139. Nosta, John (18 December 2023). "Should Artificial Intelligence Have Rights?" (https://www.psychologytoday.com/us/blog/the-digital-self/202312/should-artificial-intelligence-have-rights). *Psychology Today*. Retrieved 5 September 2024.
140. Akst, Daniel (10 April 2023). "Should Robots With Artificial Intelligence Have Moral or Legal Rights?" (https://www.wsj.com/articles/robots-ai-legal-rights-3c47ef40). *The Wall Street Journal*.
141. "Artificial General Intelligence – Do[es] the cost outweigh benefits?" (https://coe-dsai.nasscom.in/artificial-general-intelligence-do-the-cost-outweigh-benefits/). 23 August 2021. Retrieved 7 June 2023.
142. "How we can Benefit from Advancing Artificial General Intelligence (AGI) – Unite.AI" (https://www.unite.ai/artificial-general-intelligence-agi/). *www.unite.ai*. 7 April 2020. Retrieved 7 June 2023.
143. Talty, Jules; Julien, Stephan. "What Will Our Society Look Like When Artificial Intelligence Is Everywhere?" (https://www.smithsonianmag.com/innovation/artificial-intelligence-future-scenarios-180968403/). *Smithsonian Magazine*. Retrieved 7 June 2023.
144. Stevenson, Matt (8 October 2015). "Answers to Stephen Hawking's AMA are Here!" (https://www.wired.com/brandlab/2015/10/stephen-hawkings-ama/). *Wired*. ISSN 1059-1028 (https://search.worldcat.org/issn/1059-1028). Retrieved 8 June 2023.
145. Bostrom, Nick (2017). "§ Preferred order of arrival". *Superintelligence: paths, dangers, strategies* (Reprinted with corrections 2017 ed.). Oxford, United Kingdom; New York, New York, USA: Oxford University Press. ISBN 978-0-1996-7811-2.

146. Piper, Kelsey (19 November 2018). "How technological progress is making it likelier than ever that humans will destroy ourselves" (https://www.vox.com/future-perfect/2018/11/19/18097663/nick-bostrom-vulnerable-world-global-catastrophic-risks). *Vox*. Retrieved 8 June 2023.
147. Topol, Eric J.; Verghese, Abraham (2019). *Deep medicine: how artificial intelligence can make healthcare human again* (First ed.). New York, NY: Basic Books. ISBN 978-1-5416-4463-2.
148. Jumper, John; Evans, Richard; Pritzel, Alexander; Green, Tim; Figurnov, Michael; Ronneberger, Olaf; Tunyasuvunakool, Kathryn; Bates, Russ; Žídek, Augustin; Potapenko, Anna; Bridgland, Alex; Meyer, Clemens; Kohl, Simon A. A.; Ballard, Andrew J.; Cowie, Andrew (August 2021). "Highly accurate protein structure prediction with AlphaFold" (https://www.nature.com/articles/s41586-021-03819-2). *Nature*. **596** (7873): 583–589. doi:10.1038/s41586-021-03819-2 (https://doi.org/10.1038%2Fs41586-021-03819-2). ISSN 1476-4687 (https://search.worldcat.org/issn/1476-4687). PMC 8371605 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8371605).
149. Alowais, Shuroug A.; Alghamdi, Sahar S.; Alsuhebany, Nada; Alqahtani, Tariq; Alshaya, Abdulrahman I.; Almohareb, Sumaya N.; Aldairem, Atheer; Alrashed, Mohammed; Bin Saleh, Khalid; Badreldin, Hisham A.; Al Yami, Majed S.; Al Harbi, Shmeylan; Albekairy, Abdulkareem M. (22 September 2023). "Revolutionizing healthcare: the role of artificial intelligence in clinical practice" (https://bmcmededuc.biomedcentral.com/articles/10.1186/s12909-023-04698-z). *BMC Medical Education*. **23** (1): 689. doi:10.1186/s12909-023-04698-z (https://doi.org/10.1186%2Fs12909-023-04698-z). ISSN 1472-6920 (https://search.worldcat.org/issn/1472-6920). PMC 10517477 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10517477). PMID 37740191 (https://pubmed.ncbi.nlm.nih.gov/37740191).
150. Tegmark, Max (2017). *Life 3.0: being human in the age of artificial intelligence*. A Borzoi book. New York: Alfred A. Knopf. ISBN 978-1-101-94659-6.
151. Brynjolfsson, Erik; McAfee, Andrew (2016). *The second machine age: work, progress, and prosperity in a time of brilliant technologies* (First published as a Norton paperback ed.). New York London: W. W. Norton & Company. ISBN 978-0-393-35064-7.
152. Bostrom, Nick (2017). *Superintelligence: paths, dangers, strategies* (Reprinted with corrections ed.). Oxford: Oxford University Press. ISBN 978-0-19-873983-8.
153. Crawford, Kate (2021). *Atlas of AI: power, politics, and the planetary costs of artificial intelligence*. New Haven: Yale University Press. ISBN 978-0-300-20957-0.
154. Doherty, Ben (17 May 2018). "Climate change an 'existential security risk' to Australia, Senate inquiry says" (https://www.theguardian.com/environment/2018/may/18/climate-change-an-existential-security-risk-to-australia-senate-inquiry-says). *The Guardian*. ISSN 0261-3077 (https://search.worldcat.org/issn/0261-3077). Retrieved 16 July 2023.
155. MacAskill, William (2022). *What we owe the future*. New York, NY: Basic Books. ISBN 978-1-5416-1862-6.
156. Ord, Toby (2020). "Chapter 5: Future Risks, Unaligned Artificial Intelligence". *The Precipice: Existential Risk and the Future of Humanity*. Bloomsbury Publishing. ISBN 978-1-5266-0021-9.
157. Al-Sibai, Noor (13 February 2022). "OpenAI Chief Scientist Says Advanced AI May Already Be Conscious" (https://futurism.com/the-byte/openai-already-sentient). *Futurism*. Retrieved 24 December 2023.
158. Samuelsson, Paul Conrad (2019). "Artificial Consciousness: Our Greatest Ethical Challenge" (https://philosophynow.org/issues/132/Artificial_Consciousness_Our_Greatest_Ethical_Challenge). *Philosophy Now*. Retrieved 23 December 2023.
159. Kateman, Brian (24 July 2023). "AI Should Be Terrified of Humans" (https://time.com/6296234/ai-should-be-terrified-of-humans/). *TIME*. Retrieved 23 December 2023.

160. Roose, Kevin (30 May 2023). "A.I. Poses 'Risk of Extinction,' Industry Leaders Warn" (https://www.nytimes.com/2023/05/30/technology/ai-threat-warning.html). *The New York Times*. ISSN 0362-4331 (https://search.worldcat.org/issn/0362-4331). Retrieved 24 December 2023.
161. "Statement on AI Risk" (https://www.safe.ai/statement-on-ai-risk). *Center for AI Safety*. 30 May 2023. Retrieved 8 June 2023.
162. "Stephen Hawking: 'Transcendence looks at the implications of artificial intelligence – but are we taking AI seriously enough?' " (https://www.independent.co.uk/news/science/stephen-hawking-transcendence-looks-at-the-implications-of-artificial-intelligence--but-are-we-taking-ai-seriously-enough-9313474.html). *The Independent (UK)*. Archived (https://web.archive.org/web/20150925153716/http://www.independent.co.uk/news/science/stephen-hawking-transcendence-looks-at-the-implications-of-artificial-intelligence--but-are-we-taking-ai-seriously-enough-9313474.html) from the original on 25 September 2015. Retrieved 3 December 2014.
163. Herger, Mario. "The Gorilla Problem – Enterprise Garage" (https://www.enterprisegarage.io/2019/10/the-gorilla-problem/). Retrieved 7 June 2023.
164. "The fascinating Facebook debate between Yann LeCun, Stuart Russel and Yoshua Bengio about the risks of strong AI" (https://www.parlonsfutur.com/blog/the-fascinating-facebook-debate-between-yann-lecun-stuart-russel-and-yoshua). *The fascinating Facebook debate between Yann LeCun, Stuart Russel and Yoshua Bengio about the risks of strong AI* (in French). Retrieved 8 June 2023.
165. "Will Artificial Intelligence Doom The Human Race Within The Next 100 Years?" (https://www.huffpost.com/entry/artificial-intelligence-oxford_n_5689858). *HuffPost*. 22 August 2014. Retrieved 8 June 2023.
166. 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://search.worldcat.org/issn/0031-8949).
167. Bostrom, Nick (2014). *Superintelligence: Paths, Dangers, Strategies* (First ed.). Oxford University Press. ISBN 978-0-1996-7811-2.
168. Chow, Andrew R.; Perrigo, Billy (16 February 2023). "The AI Arms Race Is On. Start Worrying" (https://time.com/6255952/ai-impact-chatgpt-microsoft-google/). *TIME*. Retrieved 24 December 2023.
169. Tetlow, Gemma (12 January 2017). "AI arms race risks spiralling out of control, report warns" (https://www.ft.com/content/b56d57e8-d822-11e6-944b-e7eb37a6aa8e). *Financial Times*. Archived (https://archive.today/20220411043213/https://www.ft.com/content/b56d57e8-d822-11e6-944b-e7eb37a6aa8e) from the original on 11 April 2022. Retrieved 24 December 2023.
170. Milmo, Dan; Stacey, Kiran (25 September 2023). "Experts disagree over threat posed but artificial intelligence cannot be ignored" (https://www.theguardian.com/technology/2023/sep/25/experts-disagree-over-threat-posed-but-artificial-intelligence-cannot-be-ignored-ai). *The Guardian*. ISSN 0261-3077 (https://search.worldcat.org/issn/0261-3077). Retrieved 24 December 2023.
171. "Humanity, Security & AI, Oh My! (with Ian Bremmer & Shuman Ghosemajumder)" (https://cafe.com/stay-tuned/humanity-security-ai-oh-my-with-ian-bremmer-shuman-ghosemajumder/). *CAFE*. 20 July 2023. Retrieved 15 September 2023.
172. Hamblin, James (9 May 2014). "But What Would the End of Humanity Mean for Me?" (https://www.theatlantic.com/health/archive/2014/05/but-what-does-the-end-of-humanity-mean-for-me/361931/). *The Atlantic*. Archived (https://web.archive.org/web/20140604211145/http://www.theatlantic.com/health/archive/2014/05/but-what-does-the-end-of-humanity-mean-for-me/361931/) from the original on 4 June 2014. Retrieved 12 December 2015.
173. Titcomb, James (30 October 2023). "Big Tech is stoking fears over AI, warn scientists" (https://www.telegraph.co.uk/business/2023/10/30/big-tech-stoking-fears-over-ai-warn-scientists/). *The Telegraph*. Retrieved 7 December 2023.

174. Davidson, John (30 October 2023). "Google Brain founder says big tech is lying about AI extinction danger" (https://www.afr.com/technology/google-brain-founder-says-big-tech-is-lying-about-ai-human-extinction-danger-20231027-p5efnz). *Australian Financial Review*. Archived (https://web.archive.org/web/20231207203025/https://www.afr.com/technology/google-brain-founder-says-big-tech-is-lying-about-ai-human-extinction-danger-20231027-p5efnz) from the original on 7 December 2023. Retrieved 7 December 2023.
175. Eloundou, Tyna; Manning, Sam; Mishkin, Pamela; Rock, Daniel (17 March 2023). "GPTs are GPTs: An early look at the labor market impact potential of large language models" (https://openai.com/research/gpts-are-gpts). *OpenAI*. Retrieved 7 June 2023.
176. Hurst, Luke (23 March 2023). "OpenAI says 80% of workers could see their jobs impacted by AI. These are the jobs most affected" (https://www.euronews.com/next/2023/03/23/openai-says-80-of-workers-could-see-their-jobs-impacted-by-ai-these-are-the-jobs-most-affe). *euronews*. Retrieved 8 June 2023.
177. Sheffey, Ayelet (20 August 2021). "Elon Musk says we need universal basic income because 'in the future, physical work will be a choice' " (https://www.businessinsider.com/elon-musk-universal-basic-income-physical-work-choice-2021-8). *Business Insider*. Archived (https://web.archive.org/web/20230709081853/https://www.businessinsider.com/elon-musk-universal-basic-income-physical-work-choice-2021-8) from the original on 9 July 2023. Retrieved 8 June 2023.

Sources

UNESCO Science Report: the Race Against Time for Smarter Development (https://unesdoc.unesco.org/ark:/48223/pf0000377433/PDF/377433eng.pdf.multi). Paris: UNESCO. 11 June 2021. ISBN 978-9-2310-0450-6. Archived (https://web.archive.org/web/20220618233752/https://unesdoc.unesco.org/ark:/48223/pf0000377433/PDF/377433eng.pdf.multi) from the original on 18 June 2022. Retrieved 22 September 2021.

Chalmers, David (1996), *The Conscious Mind*, Oxford University Press.

Clocksin, William (August 2003), "Artificial intelligence and the future", *Philosophical Transactions of the Royal Society A*, vol. 361, no. 1809, pp. 1721–1748, Bibcode:2003RSPTA.361.1721C (https://ui.adsabs.harvard.edu/abs/2003RSPTA.361.1721C), doi:10.1098/rsta.2003.1232 (https://doi.org/10.1098%2Frsta.2003.1232), PMID 12952683 (https://pubmed.ncbi.nlm.nih.gov/12952683), S2CID 31032007 (https://api.semanticscholar.org/CorpusID:31032007)

Crevier, Daniel (1993). *AI: The Tumultuous Search for Artificial Intelligence*. New York, NY: BasicBooks. ISBN 0-465-02997-3.

Darrach, Brad (20 November 1970), "Meet Shakey, the First Electronic Person", *Life Magazine*, pp. 58–68

Drachman, D. (2005), "Do we have brain to spare?", *Neurology*, **64** (12): 2004–2005, doi:10.1212/01.WNL.0000166914.38327.BB (https://doi.org/10.1212%2F01.WNL.0000166914.38327.BB), PMID 15985565 (https://pubmed.ncbi.nlm.nih.gov/15985565), S2CID 38482114 (https://api.semanticscholar.org/CorpusID:38482114)

Feigenbaum, Edward A.; McCorduck, Pamela (1983), *The Fifth Generation: Artificial Intelligence and Japan's Computer Challenge to the World*, Michael Joseph, ISBN 978-0-7181-2401-4

Goertzel, Ben; Pennachin, Cassio, eds. (2006), *Artificial General Intelligence* (https://web.archive.org/web/20130320184603/http://people.inf.elte.hu/csizsekp/ai/books/artificial-general-intelligence-cognitive-technologies.9783540237334.27156.pdf) (PDF), Springer, ISBN 978-3-5402-3733-4, archived from the original (http://people.inf.elte.hu/csizsekp/ai/books/artificial-general-intelligence-cognitive-technologies.9783540237334.27156.pdf) (PDF) on 20 March 2013

Goertzel, Ben (December 2007), "Human-level artificial general intelligence and the possibility of a technological singularity: a reaction to Ray Kurzweil's *The Singularity Is Near*, and McDermott's critique of Kurzweil" (https://scholar.google.com/scholar?hl=sv&lr=&cluster=15189798216526465792), *Artificial Intelligence*, vol. 171, no. 18, Special Review Issue, pp. 1161–1173, doi:10.1016/j.artint.2007.10.011 (https://doi.org/10.1016%2Fj.artint.2007.10.011), archived (https://web.archive.org/web/20160107042341/http://scholar.google.com/scholar?hl=sv&lr=&cluster=15189798216526465792) from the original on 7 January 2016, retrieved 1 April 2009

Gubrud, Mark (November 1997), "Nanotechnology and International Security" (http://www.foresight.org/Conferences/MNT05/Papers/Gubrud/), *Fifth Foresight Conference on Molecular Nanotechnology*, archived (https://web.archive.org/web/20110529215447/http://www.foresight.org/Conferences/MNT05/Papers/Gubrud/) from the original on 29 May 2011, retrieved 7 May 2011

Howe, J. (November 1994), *Artificial Intelligence at Edinburgh University: a Perspective* (http://www.dai.ed.ac.uk/AI_at_Edinburgh_perspective.html), archived (https://web.archive.org/web/20070817012000/http://www.dai.ed.ac.uk/AI_at_Edinburgh_perspective.html) from the original on 17 August 2007, retrieved 30 August 2007

Johnson, Mark (1987), *The body in the mind*, Chicago, ISBN 978-0-2264-0317-5

Kurzweil, Ray (2005), *The Singularity is Near*, Viking Press

Lighthill, Professor Sir James (1973), "Artificial Intelligence: A General Survey", *Artificial Intelligence: a paper symposium*, Science Research Council

Luger, George; Stubblefield, William (2004), *Artificial Intelligence: Structures and Strategies for Complex Problem Solving* (https://archive.org/details/artificialintell0000luge/page/720) (5th ed.), The Benjamin/Cummings Publishing Company, Inc., p. 720 (https://archive.org/details/artificialintell0000luge/page/720), ISBN 978-0-8053-4780-7

McCarthy, John (2007b). *What is Artificial Intelligence?* (http://www-formal.stanford.edu/jmc/whatisai/whatisai.html). Stanford University. "The ultimate effort is to make computer programs that can solve problems and achieve goals in the world as well as humans."

Moravec, Hans (1988), *Mind Children*, Harvard University Press

Moravec, Hans (1998), "When will computer hardware match the human brain?" (https://web.archive.org/web/20060615031852/http://transhumanist.com/volume1/moravec.htm), *Journal of Evolution and Technology*, vol. 1, archived from the original (https://www.transhumanist.com/volume1/moravec.htm) on 15 June 2006, retrieved 23 June 2006

Nagel (1974), "What Is it Like to Be a Bat" (http://organizations.utep.edu/Portals/1475/nagel_bat.pdf) (PDF), *Philosophical Review*, **83** (4): 435–50, doi:10.2307/2183914 (https://doi.org/10.2307%2F2183914), JSTOR 2183914 (https://www.jstor.org/stable/2183914), archived (https://web.archive.org/web/20111016181405/http://organizations.utep.edu/Portals/1475/nagel_bat.pdf) (PDF) from the original on 16 October 2011, retrieved 7 November 2009

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 (1998), *Artificial Intelligence: A New Synthesis*, Morgan Kaufmann Publishers, ISBN 978-1-5586-0467-4

NRC (1999), "Developments in Artificial Intelligence" (http://www.nap.edu/readingroom/books/far/ch9.html), *Funding a Revolution: Government Support for Computing Research*, National Academy Press, archived (https://web.archive.org/web/20080112001018/http://www.nap.edu/readingroom/books/far/ch9.html) from the original on 12 January 2008, retrieved 29 September 2007

Poole, David; Mackworth, Alan; Goebel, Randy (1998), *Computational Intelligence: A Logical Approach* (http://www.cs.ubc.ca/spider/poole/ci.html), New York: Oxford University Press, archived (https://web.archive.org/web/20090725025030/http://www.cs.ubc.ca/spider/poole/ci.html) from the original on 25 July 2009, retrieved 6 December 2007

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

Sandberg, Anders; Boström, Nick (2008), *Whole Brain Emulation: A Roadmap* (http://www.fhi.ox.ac.uk/Reports/2008-3.pdf) (PDF), Technical Report #2008-3, Future of Humanity Institute, Oxford University, archived (https://web.archive.org/web/20200325021252/https://www.fhi.ox.ac.uk/reports/2008-3.pdf) (PDF) from the original on 25 March 2020, retrieved 5 April 2009

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), S2CID 55303721 (https://api.semanticscholar.org/CorpusID:55303721), 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 3 September 2020

Simon, H. A. (1965), *The Shape of Automation for Men and Management*, New York: Harper & Row

Turing, Alan (October 1950). "Computing Machinery and Intelligence" (https://academic.oup.com/mind/article/LIX/236/433/986238). *Mind*. **59** (236): 433–460. doi:10.1093/mind/LIX.236.433 (https://doi.org/10.1093%2Fmind%2FLIX.236.433). ISSN 1460-2113 (https://search.worldcat.org/issn/1460-2113). JSTOR 2251299 (https://www.jstor.org/stable/2251299). S2CID 14636783 (https://api.semanticscholar.org/CorpusID:14636783).

de Vega, Manuel; Glenberg, Arthur; Graesser, Arthur, eds. (2008), *Symbols and Embodiment: Debates on meaning and cognition*, Oxford University Press, ISBN 978-0-1992-1727-4

Wang, Pei; Goertzel, Ben (2007). "Introduction: Aspects of Artificial General Intelligence" (https://www.researchgate.net/publication/234801154). *Advances in Artificial General Intelligence: Concepts, Architectures and Algorithms: Proceedings of the AGI Workshop 2006*. IOS Press. pp. 1–16. ISBN 978-1-5860-3758-1. Archived (https://web.archive.org/web/20210218035513/https://www.researchgate.net/publication/234801154_Introduction_Aspects_of_Artificial_General_Intelligence) from the original on 18 February 2021. Retrieved 13 December 2020 – via ResearchGate.

Further reading

Aleksander, Igor (1996), *Impossible Minds* (https://archive.org/details/impossiblemindsm0000alek), World Scientific Publishing Company, ISBN 978-1-8609-4036-1

Azevedo FA, Carvalho LR, Grinberg LT, Farfel J, et al. (April 2009), "Equal numbers of neuronal and nonneuronal cells make the human brain an isometrically scaled-up primate brain" (https://www.researchgate.net/publication/24024444), *The Journal of Comparative Neurology*, **513** (5): 532–541, doi:10.1002/cne.21974 (https://doi.org/10.1002%2Fcne.21974), PMID 19226510 (https://pubmed.ncbi.nlm.nih.gov/19226510), S2CID 5200449 (https://api.semanticscholar.org/CorpusID:5200449), archived (https://web.archive.org/web/20210218035513/https://www.researchgate.net/publication/24024444_Equal_Numbers_of_Neuronal_and_Nonneuronal_Cells_Make_the_Human_Brain_an_Isometrically_Scaled-Up_Primate_Brain) from the original on 18 February 2021, retrieved 4 September 2013 – via ResearchGate

Berglas, Anthony (January 2012) [2008], *Artificial Intelligence Will Kill Our Grandchildren (Singularity)* (http://berglas.org/Articles/AIKillGrandchildren/AIKillGrandchildren.html), archived (https://web.archive.org/web/20140723053223/http://berglas.org/Articles/AIKillGrandchildren/AIKillGrandchildren.html) from the original on 23 July 2014, retrieved 31 August 2012

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.)

Gelernter, David, *Dream-logic, the Internet and Artificial Thought* (https://web.archive.org/web/20100726055120/http://www.edge.org/3rd_culture/gelernter10.1/gelernter10.1_index.html), Edge, archived from the original (http://www.edge.org/3rd_culture/gelernter10.1/gelernter10.1_index.html) on 26 July 2010, retrieved 25 July 2010

Gleick, James, "The Fate of Free Will" (review of Kevin J. Mitchell, *Free Agents: How Evolution Gave Us Free Will*, Princeton University Press, 2023, 333 pp.), *The New York Review of Books*, vol. LXXI, no. 1 (18 January 2024), pp. 27–28, 30. "Agency is what distinguishes us from machines. For biological creatures, reason and purpose come from acting in the world and experiencing the consequences. Artificial intelligences – disembodied, strangers to blood, sweat, and tears – have no occasion for that." (p. 30.)

Halal, William E. "TechCast Article Series: The Automation of Thought" (https://web.archive.org/web/20130606101835/http://www.techcast.org/Upload/PDFs/633615794236495345_TCTheAutomationofThought.pdf) (PDF). Archived from the original (http://www.techcast.org/Upload/PDFs/633615794236495345_TCTheAutomationofThought.pdf) (PDF) on 6 June 2013.

Halpern, Sue, "The Coming Tech Autocracy" (review of Verity Harding, *AI Needs You: How We Can Change AI's Future and Save Our Own*, Princeton University Press, 274 pp.; Gary Marcus, *Taming Silicon Valley: How We Can Ensure That AI Works for Us*, MIT Press, 235 pp.; Daniela Rus and Gregory Mone, *The Mind's Mirror: Risk and Reward in the Age of AI*, Norton, 280 pp.; Madhumita Murgia, *Code Dependent: Living in the Shadow of AI*, Henry Holt, 311 pp.), *The New York Review of Books*, vol. LXXI, no. 17 (7 November 2024), pp. 44–46. "'We can't realistically expect that those who hope to get rich from AI are going to have the interests of the rest of us close at heart,' ... writes [Gary Marcus]. 'We can't count on governments driven by campaign finance contributions [from tech companies] to push back.'... Marcus details the demands that citizens should make of their governments and the tech companies. They include transparency on how AI systems work; compensation for individuals if their data [are] used to train LLMs (large language model)s and the right to consent to this use; and the ability to hold tech companies liable for the harms they cause by eliminating Section 230, imposing cash penalites, and passing stricter product liability laws... Marcus also suggests... that a new, AI-specific federal agency, akin to the FDA, the FCC, or the FTC, might provide the most robust oversight.... [T]he Fordham law professor Chinmayi Sharma... suggests... establish[ing] a professional licensing regime for engineers that would function in a similar way to medical licenses, malpractice suits, and the Hippocratic oath in medicine. 'What if, like doctors,' she asks..., 'AI engineers also vowed to do no harm?'" (p. 46.)

Holte, R. C.; Choueiry, B. Y. (2003), "Abstraction and reformulation in artificial intelligence", *Philosophical Transactions of the Royal Society B*, vol. 358, no. 1435, pp. 1197–1204, doi:10.1098/rstb.2003.1317 (https://doi.org/10.1098%2Frstb.2003.1317), PMC 1693218 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1693218), PMID 12903653 (https://pubmed.ncbi.nlm.nih.gov/12903653)

Hughes-Castleberry, Kenna, "A Murder Mystery Puzzle: The literary puzzle *Cain's Jawbone*, which has stumped humans for decades, reveals the limitations of natural-language-processing algorithms", *Scientific American*, vol. 329, no. 4 (November 2023), pp. 81–82. "This murder mystery competition has revealed that although NLP (natural-language processing) models are capable of incredible feats, their abilities are very much limited by the amount of context they receive. This [...] could cause [difficulties] for researchers who hope to use them to do things such as analyze ancient languages. In some cases, there are few historical records on long-gone civilizations to serve as training data for such a purpose." (p. 82.)

Immerwahr, Daniel, "Your Lying Eyes: People now use A.I. to generate fake videos indistinguishable from real ones. How much does it matter?", *The New Yorker*, 20 November 2023, pp. 54–59. "If by 'deepfakes' we mean realistic videos produced using artificial intelligence that actually deceive people, then they barely exist. The fakes aren't deep, and the deeps aren't fake. [...] A.I.-generated videos are not, in general, operating in our media as counterfeited evidence. Their role better resembles that of cartoons, especially smutty ones." (p. 59.)

Leffer, Lauren, "The Risks of Trusting AI: We must avoid humanizing machine-learning models used in scientific research", *Scientific American*, vol. 330, no. 6 (June 2024), pp. 80–81.

Lepore, Jill, "The Chit-Chatbot: Is talking with a machine a conversation?", *The New Yorker*, 7 October 2024, pp. 12–16.

Marcus, Gary, "Artificial Confidence: Even the newest, buzziest systems of artificial general intelligence are stymmied by the same old problems", *Scientific American*, vol. 327, no. 4 (October 2022), pp. 42–45.

McCarthy, John (October 2007), "From here to human-level AI", *Artificial Intelligence*, **171** (18): 1174–1182, doi:10.1016/j.artint.2007.10.009 (https://doi.org/10.1016%2Fj.artint.2007.10.009)

McCorduck, Pamela (2004), *Machines Who Think* (2nd ed.), Natick, Massachusetts: A. K. Peters, ISBN 1-5688-1205-1

Moravec, Hans (1976), *The Role of Raw Power in Intelligence* (https://web.archive.org/web/20160303232511/http://www.frc.ri.cmu.edu/users/hpm/project.archive/general.articles/1975/Raw.Power.html), archived from the original (http://www.frc.ri.cmu.edu/users/hpm/project.archive/general.articles/1975/Raw.Power.html) on 3 March 2016, retrieved 29 September 2007

Newell, Allen; Simon, H. A. (1963), "GPS: A Program that Simulates Human Thought", in Feigenbaum, E. A.; Feldman, J. (eds.), *Computers and Thought*, New York: McGraw-Hill

Omohundro, Steve (2008), *The Nature of Self-Improving Artificial Intelligence*, presented and distributed at the 2007 Singularity Summit, San Francisco, California

Press, Eyal, "In Front of Their Faces: Does facial-recognition technology lead police to ignore contradictory evidence?", *The New Yorker*, 20 November 2023, pp. 20–26.

Roivainen, Eka, "AI's IQ: ChatGPT aced a [standard intelligence] test but showed that intelligence cannot be measured by IQ alone", *Scientific American*, vol. 329, no. 1 (July/August 2023), p. 7. "Despite its high IQ, ChatGPT fails at tasks that require real humanlike reasoning or an understanding of the physical and social world.... ChatGPT seemed unable to reason logically and tried to rely on its vast database of... facts derived from online texts."

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.)

Sutherland, J. G. (1990), "Holographic Model of Memory, Learning, and Expression", *International Journal of Neural Systems*, vol. 1–3, pp. 256–267

Vincent, James, "Horny Robot Baby Voice: James Vincent on AI chatbots", *London Review of Books*, vol. 46, no. 19 (10 October 2024), pp. 29–32. "[AI chatbot] programs are made possible by new technologies but rely on the timelelss human tendency to anthropomorphise." (p. 29.)

Williams, R. W.; Herrup, K. (1988), "The control of neuron number", *Annual Review of Neuroscience*, **11**: 423–453, doi:10.1146/annurev.ne.11.030188.002231 (https://doi.org/10.1146%2Fannurev.ne.11.030188.002231), PMID 3284447 (https://pubmed.ncbi.nlm.nih.gov/3284447)

Yudkowsky, Eliezer (2006), "Artificial General Intelligence" (https://web.archive.org/web/20090411050423/http://www.singinst.org/upload/LOGI/LOGI.pdf) (PDF), *Annual Review of Psychology*, **49**, Springer: 585–612, doi:10.1146/annurev.psych.49.1.585 (https://doi.org/10.1146%2Fannurev.psych.49.1.585), ISBN 978-3-5402-3733-4, PMID 9496632 (https://pubmed.ncbi.nlm.nih.gov/9496632), archived from the original (http://www.singinst.org/upload/LOGI//LOGI.pdf) (PDF) on 11 April 2009

Yudkowsky, Eliezer (2008), "Artificial Intelligence as a Positive and Negative Factor in Global Risk", *Global Catastrophic Risks*, Bibcode:2008gcr..book..303Y (https://ui.adsabs.harvard.edu/abs/2008gcr..book..303Y), doi:10.1093/oso/9780198570509.003.0021 (https://doi.org/10.1093%2Foso%2F9780198570509.003.0021), ISBN 978-0-1985-7050-9

Zucker, Jean-Daniel (July 2003), "A grounded theory of abstraction in artificial intelligence", *Philosophical Transactions of the Royal Society B*, vol. 358, no. 1435, pp. 1293–1309, doi:10.1098/rstb.2003.1308 (https://doi.org/10.1098%2Frstb.2003.1308), PMC 1693211 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1693211), PMID 12903672 (https://pubmed.ncbi.nlm.nih.gov/12903672)

External links

- The AGI portal maintained by Pei Wang (https://cis.temple.edu/~pwang/AGI-Intro.html)