

Renaud Jerez
Adideath
18/01-15/03/14



WIKIPEDIA
The Free Encyclopedia

- Main page
- Contents
- Featured content
- Current events
- Random article
- Donate to Wikipedia
- Wikimedia Shop

- Interaction
 - Help
 - About Wikipedia
 - Community portal
 - Recent changes
 - Contact page

Tools

Print/export

- Languages
 - Беларуская
 - Català
 - Čeština
 - Deutsch
 - Eesti
 - Ελληνικά
 - Español
 - Français
 - Italiano
 - עברית
 - Latina
 - Latviešu
 - Lietuvių
 - Magyar
 - Nederlands
 - नेपाली
 - 日本語
 - Polski
 - Português
 - Română
 - Русский
 - Simple English
 - Slovenčina
 - Slovenščina
 - Српски / srpski
 - Srpskohrvatski / српскохрватски
 - Suomi
 - Svenska
 - தமிழ்
 - Татарча/tatarça
 - Українська
 - Tiếng Việt
 - 中文

Edit links

Create account Log in

Article Talk

Read Edit View history

Search

Technological singularity

From Wikipedia, the free encyclopedia

This article has multiple issues. Please help [improve it](#) or discuss these issues on the [talk page](#).

- The **lead section of this article may need to be rewritten.** (*November 2013*)
- This article uses **bare URLs** for citations, which may be threatened by [link rot](#). (*August 2013*)

The **technological singularity**, or simply **the singularity**, is a hypothetical moment in time when **artificial intelligence** will have progressed to the point of a **greater-than-human intelligence**, radically changing civilization, and perhaps human nature.^[1] Since the capabilities of such an intelligence may be difficult for a human to comprehend, the technological singularity is often seen as an occurrence (akin to a [gravitational singularity](#)) beyond which the future course of human history is unpredictable or even unfathomable.

The first use of the term "singularity" in this context was by mathematician [John von Neumann](#). In 1958, regarding a summary of a conversation with von Neumann, [Stanislaw Ulam](#) described "ever accelerating progress of technology and changes in the mode of human life, which gives the appearance of approaching some essential singularity in the history of the race beyond which human affairs, as we know them, could not continue".^[2] The term was popularized by science fiction writer [Vernor Vinge](#), who argues that [artificial intelligence](#), [human biological enhancement](#), or [brain-computer interfaces](#) could be possible causes of the singularity.^[3] Futurist [Ray Kurzweil](#) cited von Neumann's use of the term in a foreword to von Neumann's classic *The Computer and the Brain*.

Proponents of the singularity typically postulate an "intelligence explosion",^{[4][5]} where superintelligences design successive generations of increasingly powerful minds, that might occur very quickly and might not stop until the agent's cognitive abilities greatly surpass that of any human.

Kurzweil predicts the singularity to occur around 2045^[6] whereas Vinge predicts some time before 2030.^[7] At the 2012 [Singularity Summit](#), Stuart Armstrong did a study of artificial generalized intelligence (AGI) predictions by experts and found a wide range of predicted dates, with a median value of 2040. His own prediction on reviewing the data is that there's an 80% probability that the singularity will occur between 2017 and 2112.^[8]

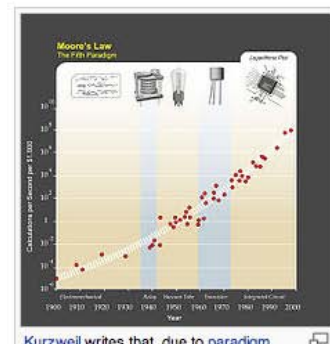
Contents [hide]

- Basic concepts
- History of the idea
- Intelligence explosion
 - Intelligence improvements
 - Impact
 - Existential risk
 - Implications for human society
- Accelerating change
- Criticisms
- In popular culture
- See also
- Notes
- References
- External links
 - Essays and articles
 - Other links

Basic concepts [edit]

Many of the most recognized writers on the singularity, such as [Vernor Vinge](#) and [Ray Kurzweil](#), define the concept in terms of the technological creation of superintelligence, and argue that it is difficult or impossible for present-day humans to predict what a post-singularity would be like, due to the difficulty of imagining the intentions and capabilities of superintelligent entities.^{[6][7][9]} The term "technological singularity" was originally coined by Vinge, who made an analogy between the breakdown in our ability to predict what would happen after the development of superintelligence and the breakdown of the predictive ability of modern physics at the [space-time singularity](#) beyond the [event horizon](#) of a black hole.^[9]

Some writers use "the singularity" in a broader way to refer to any radical changes in our society brought about by new technologies such as [molecular nanotechnology](#),^{[10][11][12]} although Vinge and other prominent writers specifically state that without superintelligence, such changes would not qualify as a true singularity.^[7] Many writers also tie the singularity to observations of exponential growth in various technologies (with [Moore's Law](#) being the most prominent example), using such observations as a basis for predicting that the singularity is likely to happen sometime within the 21st century.^{[11][13]}



A technological singularity includes the concept of an intelligence explosion, a term coined in 1965 by I. J. Good.^[14] Although technological progress has been accelerating, it has been limited by the basic intelligence of the human brain, which has not, according to Paul R. Ehrlich, changed significantly for millennia.^[15] However, with the increasing power of computers and other technologies, it might eventually be possible to build a machine that is more intelligent than humanity.^[16] If a superhuman intelligence were to be invented—either through the [amplification of human intelligence](#) or through artificial intelligence—it would bring to bear greater problem-solving and inventive skills than current humans are capable of. It could then design an even more capable machine, or re-write its own source code to become even more intelligent. This more capable machine could then go on to design a machine of yet greater capability. These iterations of [recursive self-improvement](#) could accelerate, potentially allowing enormous qualitative change before any upper limits imposed by the laws of physics or theoretical computation set in.^{[17][18][19]}

shifts, a trend of exponential growth extends Moore's law to integrated circuits from earlier transistors, vacuum tubes, relays, and electromechanical computers. He predicts that the exponential growth will continue, and that in a few decades the computing power of all computers will exceed that of human brains, with superhuman artificial intelligence appearing around the same time.

The exponential growth in computing technology suggested by Moore's Law is commonly cited as a reason to expect a singularity in the relatively near future, and a number of authors have proposed generalizations of Moore's Law. Computer scientist and futurist Hans Moravec proposed in a 1998 book^{[20][citation needed]} that the exponential growth curve could be extended back through earlier computing technologies prior to the [integrated circuit](#). Futurist Ray Kurzweil postulates a [law of accelerating returns](#) in which the speed of technological change (and more generally, all evolutionary processes^[21]) increases exponentially, generalizing Moore's Law in the same manner as Moravec's proposal, and also including material technology (especially as applied to [nanotechnology](#)), medical technology and others.^[22] Between 1986 and 2007, machines' application-specific capacity to compute information per capita has roughly doubled every 14 months; the per capita capacity of the world's general-purpose computers has doubled every 18 months; the global telecommunication capacity per capita doubled every 34 months; and the world's storage capacity per capita doubled every 40 months.^[23] Like other authors, though, Kurzweil reserves the term "singularity" for a rapid increase in intelligence (as opposed to other technologies), writing for example that "The Singularity will allow us to transcend these limitations of our biological bodies and brains ... There will be no distinction, post-Singularity, between human and machine".^[24] He believes that the "design of the human brain, while not simple, is nonetheless a billion times simpler than it appears, due to massive redundancy".^[25] According to Kurzweil, the reason why the brain has a messy and unpredictable quality is because the brain, like most biological systems, is a "probabilistic fractal".^[26] He also defines his predicted date of the singularity (2045) in terms of when he expects computer-based intelligences to significantly exceed the sum total of human brainpower, writing that advances in computing before that date "will not represent the Singularity" because they do "not yet correspond to a profound expansion of our intelligence."^[27]

The term "technological singularity" reflects the idea that such change may happen suddenly, and that it is difficult to predict how such a new world would operate.^{[28][29]} It is unclear whether an intelligence explosion of this kind would be beneficial or harmful, or even an [existential threat](#),^{[30][31]} as the issue has not been dealt with by most [artificial general intelligence](#) researchers, although the topic of [friendly artificial intelligence](#) is investigated by the [Future of Humanity Institute](#) and the Singularity Institute for Artificial Intelligence, which is now the [Machine Intelligence Research Institute](#).^[28]

Many prominent technologists and academics dispute the plausibility of a technological singularity, including Jeff Hawkins, John Holland, Jaron Lanier, and Gordon Moore, whose Moore's Law is often cited in support of the concept.^{[32][33]}

History of the idea [\[edit\]](#)

In 1847, R. Thornton, the editor of *The Expounder of Primitive Christianity*,^[34] wrote about the recent invention of a four-function [mechanical calculator](#):

...such machines, by which the scholar may, by turning a crank, grind out the solution of a problem without the fatigue of mental application, would by its introduction into schools, do incalculable injury. But who knows that such machines when brought to greater perfection, may not think of a plan to remedy all their own defects and then grind out ideas beyond the ken of mortal mind!

In 1909, the historian Henry Adams wrote an essay, *The Rule of Phase Applied to History*,^[35] in which he developed a "physical theory of history" by applying the law of inverse squares to historical periods, proposing a "Law of the Acceleration of Thought." Adams interpreted history as a process moving towards an "equilibrium", and speculated that this process would "bring Thought to the limit of its possibilities in the year 1921. It may well be!", adding that the "consequences may be as surprising as the change of water to vapor, of the worm to the butterfly, of radium to electrons."^[36] The futurist John Smart has called Adams "Earth's First Singularity Theorist".^[37]

In 1951, Alan Turing spoke of machines outstripping humans intellectually.^[38]

once the machine thinking method has started, it would not take long to outstrip our feeble powers. ... At some stage therefore we should have to expect the machines to take control, in the way that is mentioned in Samuel Butler's *Erewhon*.

In the mid fifties, Stanislaw Ulam had a conversation with John von Neumann in which von Neumann spoke of "ever accelerating progress of technology and changes in the mode of human life, which gives the appearance of approaching some essential singularity in the history of the race beyond which human affairs, as we know them, could not continue."^[2]

In 1965, I. J. Good first wrote of an "intelligence explosion", suggesting that if machines could even slightly surpass human intellect, they could improve their own designs in ways unforeseen by their designers, and thus [recursively](#) augment themselves into far greater intelligences. The first such improvements might be small, but as the machine became more intelligent it would become better at becoming more intelligent, which could lead to a cascade of self-improvements and a sudden surge to superintelligence (or a singularity).

In 1983, mathematician and author Vernor Vinge greatly popularized Good's notion of an intelligence explosion in a number of writings, first addressing the topic in print in the January 1983 issue of *Omni* magazine. In this op-ed piece, Vinge seems to have been the first to use the term "singularity" in a way that was specifically tied to the creation of intelligent machines,^{[39][40]} writing:

We will soon create intelligences greater than our own. When this happens, human history will have reached a kind of singularity, an intellectual transition as impenetrable as the knotted space-time at the center of a black hole, and the world will pass far beyond our understanding. This singularity, I believe, already haunts a number of science-fiction writers. It makes realistic extrapolation to an interstellar future impossible. To write a story set more than a century hence, one needs a nuclear war in between ... so that the world remains intelligible.

In 1984, Samuel R. Delany used "cultural fugue" as a plot device in his science-fiction novel *Stars in My Pocket Like Grains of Sand*; the terminal runaway of technological and cultural complexity in effect destroys all life on any world on which it transpires, a process poorly understood by the novel's characters, and against which they seek a stable defense. In 1985, Ray Solomonoff introduced the notion of "infinity point"^[41] in the time-scale of artificial intelligence, analyzed the magnitude of the "future shock" that "we can expect from our AI expanded scientific community" and on social effects. Estimates were made "for when these milestones would occur, followed by some suggestions for the more effective utilization of the extremely rapid technological growth that is expected".

Vinge also popularized the concept in SF novels such as *Marooned in Realtime* (1986) and *A Fire Upon the Deep* (1992). The former is set in a world of rapidly [accelerating change](#) leading to the emergence of more and more sophisticated technologies separated by shorter and shorter time-intervals, until a point beyond human comprehension is reached. The latter starts with an imaginative description of the evolution of a superintelligence passing through exponentially accelerating developmental stages ending in a [transcendent](#), almost [omnipotent](#) power unfathomable by mere humans. Vinge also implies that the development may not stop at this level.

In his 1988 book *Mind Children*, computer scientist and futurist Hans Moravec generalizes Moore's law to make predictions about the future of artificial life. Moravec outlines a timeline and a scenario in this regard,^{[42][43]} in that robots will evolve into a new series of artificial species, starting around 2030–2040.^[44] In *Robot: Mere Machine to Transcendent Mind*, published in 1998, Moravec further considers the implications of evolving [robot intelligence](#), generalizing Moore's law to technologies predating the [integrated circuit](#), and speculating about a coming "mind fire" of rapidly expanding superintelligence, similar to Vinge's ideas.

A 1993 article by Vinge, "The Coming Technological Singularity: How to Survive in the Post-Human Era",^[7] spread widely on the internet and helped to popularize the idea.^[45] This article contains the oft-quoted statement, "Within thirty years, we will have the technological means to create superhuman intelligence. Shortly after, the human era will be ended." Vinge refines his estimate of the time-scales involved, adding, "I'll be surprised if this event occurs before 2005 or after 2030."

Vinge predicted four ways the singularity could occur.^[46]

1. The development of computers that are "awake" and superhumanly intelligent
2. Large computer networks (and their associated users) may "wake up" as a superhumanly intelligent entity
3. Computer/human interfaces may become so intimate that users may reasonably be considered superhumanly intelligent
4. Biological science may find ways to improve upon the natural human intellect

Vinge continues by predicting that superhuman intelligences will be able to enhance their own minds faster than their human creators. "When greater-than-human intelligence drives progress," Vinge writes, "that progress will be much more rapid." He predicts that this [feedback loop](#) of self-improving intelligence will cause large amounts of technological progress within a short period, and states that the creation of superhuman intelligence represents a breakdown in humans' ability to model their future. His argument was that authors cannot write realistic characters who surpass the human intellect, as the thoughts of such an intellect would be beyond the ability of humans to express. Vinge named this event "the Singularity".

Damien Broderick's popular science book *The Spike* (1997) was the first^[citation needed] to investigate the technological singularity in detail.

In 2000, Bill Joy, a prominent technologist and a co-founder of Sun Microsystems, voiced concern over the potential dangers of the singularity.^[47]

In 2005, Ray Kurzweil published *The Singularity is Near*, which brought the idea of the singularity to the popular media both through the book's accessibility and through a publicity campaign that included an appearance on *The Daily Show with Jon Stewart*.^[48] The book stirred intense controversy, in part because Kurzweil's [utopian](#) predictions contrasted starkly with other, darker visions of the possibilities of the singularity.^[original research?] Kurzweil, his theories, and the controversies surrounding it were the subject of Barry Ptolemy's documentary *Transcendent Man*.

In 2007, Eliezer Yudkowsky suggested that many of the varied definitions that have been assigned to "singularity" are mutually incompatible rather than mutually supporting.^[11] For example, Kurzweil extrapolates current technological trajectories past the arrival of self-improving AI or superhuman intelligence, which Yudkowsky argues represents a tension with both I. J. Good's proposed discontinuous upswing in intelligence and Vinge's thesis on unpredictability.

In 2008, Robin Hanson (taking "singularity" to refer to sharp increases in the exponent of economic growth) listed the [Agricultural](#) and [Industrial Revolutions](#) as past singularities. Extrapolating from such past events, Hanson proposes that the next economic singularity should increase [economic growth](#) between 60 and 250 times. An innovation that allowed for the replacement of virtually all human labor could trigger this event.^[49]

In 2009, Kurzweil and X-Prize founder Peter Diamandis announced the establishment of [Singularity University](#), whose stated mission is "to assemble, educate and inspire a cadre of leaders who strive to understand and facilitate the development of exponentially advancing technologies in order to address humanity's grand challenges."^[50] Funded by [Google](#), [Autodesk](#), [ePlanet Ventures](#), and a group of [technology-industry](#) leaders, Singularity University is based at [NASA's Ames Research Center](#) in Mountain View, California. The not-for-profit organization runs an annual ten-week graduate program during the northern-hemisphere summer that covers ten different technology and allied tracks, and a series of executive programs throughout the year.

In 2010, Aubrey de Grey applied the term the "Methuselarity"^[51] to the point at which medical technology improves so fast that [expected human lifespan](#) increases by more than one year per year. In "Apocalyptic AI – Visions of Heaven in Robotics, Artificial Intelligence, and Virtual Reality"^[52] (2010), Robert Geraci offers an account of the developing "cyber-theology" inspired by Singularity studies. The 1996 novel *Holy Fire* by Bruce Sterling explores some of those themes and postulates that a Methuselarity will become a [gerontocracy](#).

In 2011, Kurzweil noted existing trends and concluded that it appeared increasingly likely that the singularity would occur around 2045. He told *Time* magazine: "We will successfully reverse-engineer the human brain by the mid-2020s. By the end of that decade, computers will be capable of human-level intelligence."^[53]

Intelligence explosion [\[edit\]](#)

The notion of an "intelligence explosion" was first described thus by Good (1965), who speculated on the effects of superhuman machines:

Let an ultraintelligent machine be defined as a machine that can far surpass all the intellectual activities of any man however clever. Since the design of machines is one of these intellectual activities, an ultraintelligent machine could design even better machines; there would then unquestionably be an 'intelligence explosion,' and the intelligence of man would be left far behind. Thus the first ultraintelligent machine is the last invention that man need ever make.

Most proposed methods for creating superhuman or [transhuman](#) minds fall into one of two categories: intelligence amplification of human brains and artificial intelligence. The means speculated to produce intelligence augmentation are numerous, and include [bioengineering](#), [genetic engineering](#), [nootropic drugs](#), AI assistants, direct brain-computer interfaces and [mind uploading](#). The existence of multiple paths to an intelligence explosion makes a singularity more likely; for a singularity to not occur they would all have to fail.^[9]

Hanson (1998) is skeptical of human intelligence augmentation, writing that once one has exhausted the "low-hanging fruit" of easy methods for increasing human intelligence, further improvements will become increasingly difficult to find. Despite the numerous speculated means for amplifying human intelligence, non-human artificial intelligence (specifically [seed AI](#)) is the most popular option for organizations trying to advance the singularity.^[citation needed]

Whether or not an intelligence explosion occurs depends on three factors.^[54] The first, accelerating factor, is the new intelligence enhancements made possible by each previous improvement. Contrariwise, as the intelligences become more advanced, further advances will become more and more complicated, possibly overcoming the advantage of increased intelligence. Each improvement must be able to beget at least one more improvement, on average, for the singularity to continue. Finally the laws of physics will eventually prevent any further improvements.