

The New York Times

This copy is for your personal, noncommercial use only. You can order presentation-ready copies for distribution to your colleagues, clients or customers [here](#) or use the "Reprints" tool that appears next to any article. Visit www.nytimes.com for samples and additional information. [Order a reprint of this article now.](#)



May 24, 2009

The Coming Superbrain

By [JOHN MARKOFF](#)

Mountain View, Calif. — It's summertime and the Terminator is back. A sci-fi movie thrill ride, "[Terminator Salvation](#)" comes complete with a malevolent artificial intelligence dubbed Skynet, a military R.&D. project that gained self-awareness and concluded that humans were an irritant — perhaps a bit like athlete's foot — to be dispatched forthwith.

The notion that a self-aware computing system would emerge spontaneously from the interconnections of billions of computers and computer networks goes back in science fiction at least as far as [Arthur C. Clarke's](#) "Dial F for Frankenstein." A prescient short story that appeared in 1961, it foretold an ever-more-interconnected telephone network that spontaneously acts like a newborn baby and leads to global chaos as it takes over financial, transportation and military systems.

Today, artificial intelligence, once the preserve of science fiction writers and eccentric computer prodigies, is back in fashion and getting serious attention from [NASA](#) and from Silicon Valley companies like [Google](#) as well as a new round of start-ups that are designing everything from next-generation search engines to machines that listen or that are capable of walking around in the world. A.I.'s new respectability is turning the spotlight back on the question of where the technology might be heading and, more ominously, perhaps, whether computer intelligence will surpass our own, and how quickly.

The concept of ultrasmart computers — machines with "greater than human intelligence" — was dubbed "The Singularity" in a [1993 paper](#) by the computer scientist and science fiction writer Vernor Vinge. He argued that the acceleration of technological progress had led to "the edge of change comparable to the rise of human life on Earth." This thesis has long struck a chord here in Silicon Valley.

Artificial intelligence is already used to automate and replace some human functions with computer-driven machines. These machines can see and hear, respond to questions, learn, draw inferences and solve problems. But for the Singulatarians, A.I. refers to machines that will be both self-aware and superhuman in their intelligence, and capable of designing better computers and robots faster than humans can today. Such a shift, they say, would lead to a vast acceleration in technological improvements of all kinds.

The idea is not just the province of science fiction authors; a generation of computer hackers, engineers and programmers have come to believe deeply in the idea of exponential technological change as explained by Gordon Moore, a co-founder of the chip maker [Intel](#).

In 1965, Dr. Moore first described the repeated doubling of the number transistors on silicon chips with each new technology generation, which led to an acceleration in the power of computing. Since then "Moore's

Law” — which is not a law of physics, but rather a description of the rate of industrial change — has come to personify an industry that lives on Internet time, where the Next Big Thing is always just around the corner.

Several years ago the artificial-intelligence pioneer Raymond Kurzweil took the idea one step further in his 2005 book, [“The Singularity Is Near: When Humans Transcend Biology.”](#) He sought to expand Moore’s Law to encompass more than just processing power and to simultaneously predict with great precision the arrival of post-human evolution, which he said would occur in 2045.

In Dr. Kurzweil’s telling, rapidly increasing computing power in concert with cyborg humans would then reach a point when machine intelligence not only surpassed human intelligence but took over the process of technological invention, with unpredictable consequences.

Profiled in the documentary [“Transcendent Man.”](#) which had its premier last month at the TriBeCa Film Festival, and with his own Singularity movie due later this year, Dr. Kurzweil has become a one-man marketing machine for the concept of post-humanism. He is the co-founder of [Singularity University](#), a school supported by Google that will open in June with a grand goal — to “assemble, educate and inspire a cadre of leaders who strive to understand and facilitate the development of exponentially advancing technologies and apply, focus and guide these tools to address humanity’s grand challenges.”

Not content with the development of superhuman machines, Dr. Kurzweil envisions “uploading,” or the idea that the contents of our brain and thought processes can somehow be translated into a computing environment, making a form of immortality possible — within his lifetime.

That has led to no shortage of raised eyebrows among hard-nosed technologists in the engineering culture here, some of whom describe the Kurzweilian romance with supermachines as a new form of religion.

The science fiction author Ken MacLeod described the idea of the singularity as “the Rapture of the nerds.” Kevin Kelly, an editor at Wired magazine, notes, “People who predict a very utopian future always predict that it is going to happen before they die.”

However, Mr. Kelly himself has not refrained from speculating on where communications and computing technology is heading. He is at work on his own book, “The Technium,” forecasting the emergence of a global brain — the idea that the planet’s interconnected computers might someday act in a coordinated fashion and perhaps exhibit intelligence. He just isn’t certain about how soon an intelligent global brain will arrive.

Others who have observed the increasing power of computing technology are even less sanguine about the future outcome. The computer designer and venture capitalist William Joy, for example, wrote a [pessimistic essay in Wired](#) in 2000 that argued that humans are more likely to destroy themselves with their technology than create a utopia assisted by superintelligent machines.

Mr. Joy, a co-founder of [Sun Microsystems](#), still believes that. “I wasn’t saying we would be supplanted by something,” he said. “I think a catastrophe is more likely.”

Moreover, there is a hot debate here over whether such machines might be the “machines of loving grace,” of the Richard Brautigan poem, or something far darker, of the “Terminator” ilk.

“I see the debate over whether we should build these artificial intellects as becoming the dominant political

question of the century,” said Hugo de Garis, an Australian artificial-intelligence researcher, who has written a book, “The Artilect War,” that argues that the debate is likely to end in global war.

Concerned about the same potential outcome, the A.I. researcher Eliezer S. Yudkowsky, an employee of the Singularity Institute, has proposed the idea of “friendly artificial intelligence,” an engineering discipline that would seek to ensure that future machines would remain our servants or equals rather than our masters.

Nevertheless, this generation of humans, at least, is perhaps unlikely to need to rush to the barricades. The artificial-intelligence industry has advanced in fits and starts over the past half-century, since the term “artificial intelligence” was coined by the [Stanford University](#) computer scientist John McCarthy in 1956. In 1964, when Mr. McCarthy established the Stanford Artificial Intelligence Laboratory, the researchers informed their Pentagon backers that the construction of an artificially intelligent machine would take about a decade. Two decades later, in 1984, that original optimism hit a rough patch, leading to the collapse of a crop of A.I. start-up companies in Silicon Valley, a time known as “the A.I. winter.”

Such reversals have led the veteran Silicon Valley technology forecaster Paul Saffo to proclaim: “never mistake a clear view for a short distance.”

Indeed, despite this high-technology heartland’s deeply held consensus about exponential progress, the worst fate of all for the Valley’s digerati would be to be the generation before the generation that lives to see the singularity.

“Kurzweil will probably die, along with the rest of us not too long before the ‘great dawn,’ ” said Gary Bradski, a Silicon Valley roboticist. “Life’s not fair.”

[Copyright 2009 The New York Times Company](#)

[Privacy Policy](#) | [Search](#) | [Corrections](#) | [RSS](#) | [First Look](#) | [Help](#) | [Contact Us](#) | [Work for Us](#) | [Site Map](#)