# Institute of Futurology
## Knowing our future

## The improbable scenario of the runaway future:
## deconstructing the Singularity
(originally published as a blog for the World Future Society)

*By Michael Lee*

"It is sometimes worthwhile to take science fiction seriously."
Irving Good *Speculations Concerning the First Ultra-intelligent Machine* (1965)

Recently, it was time to celebrate what technology can do in medicine when a patient at the Ohio State University Wexner Medical Center, 23 year-old Ian Burkhart, who was paralyzed a few years ago in a swimming accident, was able to use his thoughts to move his hand again through the Neurobridge system. This technology deployed a microchip implanted into his brain, with 96 electrodes for "reading" some of his thoughts it could recognize, like a pre-programmed code. Then his command to move his hand was sent from these electrodes, via a computer, to receiver-electrodes attached to his sleeve which stimulated his muscles.[1] Empowering a patient in this way to regain some control over paralyzed muscles and limbs must have been incredibly liberating.

Outside of health and medicine, though, enhancing human brains with chip implants could open up a Pandora's box of unintended consequences, including permanently fracturing the human race into technologically enhanced "haves" and unenhanced "have nots", compounding the world's existing dangerous divisions along class, racial and ideological lines. Such an unintended consequence would be a very real possibility given the intensely competitive nature of human beings whom evolution has honed, over millennia, to seek every competitive advantage for survival.

The issue of intelligence has always been a key to our survival and is a core concept in the kind of future forecast by advocates of the Singularity. Yet its true nature is not well understood, especially in discussions about human versus machine intelligence.

---

[1] Rosa Prince, The Telegraph | June 25, 2014 | Last Updated: Jun 25 3:19 PM ET
http://news.nationalpost.com/2014/06/25/science-fiction-come-true-paralyzed-man-moves-arm-with-his-thoughts-using-microchip-in-his-brain/

© Institute of Futurology, July 2014

The Singularity is an imagined future event in which runaway change, caused by an intelligence explosion and/or a hyper-expansion of technology, and driven to exponential levels through the mechanism of a feedback loop between increases in intelligence capacity, efficiency and technological development, alters human society beyond what its current mental models can conceive, or handle.

Upon close examination, however, this runaway future, posited by such thinkers as visionary technology futurist Ray Kurzweil, Vernor Vinge and Irving Good, turns out to be an improbable scenario. Some of its theory is pure speculation, unencumbered by the inconvenient truths of natural limits, economic and technology lifecycles, the law of causation or historical patterns of social evolution.

In fact, there's more evidence that the very opposite could happen, namely, a gradual slowing down of human and technological progress caused by natural limits, including environmental factors and energy resources, as well as by alarming trends towards human depopulation and ageing of whole global populations.

Natural limits to economic growth received serious international attention when the Club of Rome published their landmark 1972 report *The Limits to Growth*. This analysis of a coming global economic slowdown, which stood in stark contrast to the heady optimism of *The Year 2000* forecasts by Herman Kahn and Antony Weiner in 1967, was given an extensive 30-year update in 2004, showing that time had proven many of its negative predictions to have been robust and sound.

Fortunately, as I will argue in this essay, there's a strong chance that humanity can steer its ship of civilization safely between the twin hazards of a Singularity or a global slowdown scenario.

Now, let's examine the case for a Runaway Future.

Modern prophecies about a coming Singularity really got going in 1965 when a paper by Irving Good called "Speculations Concerning the First Ultra-intelligent Machine" appeared in the journal *Advances in Computers*. Good based his essay on talks he'd given at a conference on Biocommunications at the Neuropsychiatric Institute at University of California in 1962 and at sessions on Artificial Intelligence at the Institute of Electrical and Electronics Engineers in 1963.

In this paper, he wrote: "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." He also stated "A man cannot learn more than ten million statements in a lifetime. A machine could already store this amount of information without much difficulty…"

From its inception, this futuristic discussion about the arrival of ultra-intelligence on earth has been bedevilled by a failure to distinguish between sheer computational power and the extraordinary richness and multi-dimensionality of human thought and intelligence. Good's point that computers can store more information than humans is disingenuous. So what? It's not the number of words in *War and Peace* that makes Tolstoy's novel a great work of art but its sustained imaginative synthesis of form and content in a breathtaking recreation of life in Russia leading up to the national trauma of Napoleon's invasion of 1812. I've never understood why multiplying the capacity to process and store information can be described as "ultra-intelligence" when it doesn't even mimic multi-dimensional human intelligence in the first place, let alone constitute some sort of mysterious new consciousness called *ultra*-intelligence.

Putting that caveat to one side for a moment, let's examine the main concepts of the Singularity to enable us to visualize how such a scenario could unfold in the real world.

Anders Sandberg's taxonomy of technological singularity[2] outlines its three main concepts:

1. Accelerating change – exponential technology development whereby smarter systems can improve themselves, producing a powerful feedback loop.
2. Prediction horizon – whereby the scale of technology change and/or the emergence of super-intelligence would make the future impossible to predict beyond the Singularity, given our current level of knowledge.
3. Intelligence explosion – artificial intelligence will increase at such a rate that it will outstrip human intelligence and lead to the emergence of superhuman intelligence which, in turn, will create a post-human future or society.

---

[2] "An Overview of Models of Technological Singularity" by Anders Sandberg, The Transhumanist Reader, p.377-8.

I'm sceptical about both an intelligence explosion and runaway technological change. Good's "ultra-intelligence" has never arrived as a consciousness or form of computer intelligence. And I'd dismiss outright the possibility that machines on their own will ever take over society, due to the one-dimensionality of machine intelligence compared to multi-dimensional human intelligence. I don't foresee robots, computers or networks of computers awakening one day to full self-awareness or consciousness. That's because consciousness emerges from the integrated physiological and mental processes of living as an organism interacting continuously with the four-dimensional space-time continuum of its environment. Since they aren't living organisms, robots and computers cannot enjoy dynamic, holistic, integrated physiological and mental interaction with 4D space-time. How then would they ever develop a full, living consciousness or become self-aware beings with their own identity, personality and values?

I also can't see technology itself ever developing to a point where humanity loses control of it in the proposed explosion of change, given that we control the tools and methods of science and as well as the governance of all systems.

Consequently, the only Singularity I can foresee having a remote chance of happening would be the rise to power of a race of technologically enhanced superhumans. Here, a merger would take place of human and machine to create superhumans who would then go on to develop a civilization that is fundamentally different from anything we've seen in history. Even in this scenario, it should be borne in mind that the interface between human intelligence and machine intelligence is a Gordian knot of entangled philosophical, ethical, legal and technical issues which may never be satisfactorily resolved in society *outside of the field of medicine*.

Whatever kind of Singularity is being envisaged, depending on where you stand in Sandberg's classification above, its advent has been generally predicted to occur between 2029-2045.

But there are four virtually insurmountable challenges which must be overcome for the technological Singularity to come about:

Challenge 1

True intelligence is much more than the power to process and store information and emerges from holistic consciousness as a force greater than the sum of human perception, memory, anticipation, creativity and reasoning; such emergent holistic consciousness, I argue, *cannot* be mimicked or reproduced by artificial means.

Challenge 2

Machines on their own *cannot* have their own organic sensory apparatus for perception and so will never possess a living, experiential consciousness.

Challenge 3

Machines *cannot* build civilisations or societies better than humans because they have no culture, philosophy, worldview, ethics, aesthetic or history.

Challenge 4

Human beings, through the mastery of science, *cannot* relinquish social governance and governance of systems; rather, they'll continue to create increasing levels of order.

As a result of these four principles, I rule out completely a Palace Revolt by Robots scenario, whereby robots and computer networks might successfully overpower their human rulers. I also rule out completely a Runaway Technology scenario whereby humans lose control of the forces of technology development and governance of technology systems, unless they *voluntarily* relinquished that role. Otherwise, how would we lose control of technologies to the point that they could control us when they were created only through the application of the science humans invented? As guardians of the laws of nature and the principles of logic, it is humans who hold the keys to knowledge and governance of our world.

On a probability scale of 0-1, where 1 is certainty and 0 represents impossibility, I would rate the Palace Revolt of Robots and the Runaway Technology scenarios right at the end of the scale at 0. Zero.

For me, the most probable Singularity would be when an elite, global class of technologically enhanced, and wealthy, super-humans, or cyborgs, wrests control of the United Nations from within and begins to govern society and culture according to the principles of the philosophy of Transhumanism, implementing an absolute technocracy. This would only be possible after decades of evolutionary cultural change culminating in the take-over by cyborgs somewhere in the 2075-2100 time frame. Such a political event would then generate a complete transformation of society into something currently unimaginable, justifying the use of the term Singularity. This is definitely a real political and technological *possibility* but one which I would consider, in practical terms, to be undesirable, immensely controversial and highly unlikely. Let's call this 2075 scenario the Techno-Transhuman Transfiguration.

On the probability scale, I would rate this at the lower end of the scale at 0.2. So, yes, as outrageous as it sounds, there's perhaps a twenty percent chance that the world will be ruled by technologically enhanced superhumans in the final decades of the 21st century. After that hypothetical threshold, all bets would be off for the future of society as human control over technology would have been voluntarily handed over to human-machine systems.

Right now, though, we see that several decades after the invention of computers and robots, and just under 50 years since Irving Good predicted the rise of ultra-intelligence, machine intelligence is still woefully below the quality of human intelligence, which, as I have argued, is multi-dimensional, holistic and emergent in character. Even if intelligence is looked at in one narrow way as the ability to reason and produce insight in original ways, machine intelligence cannot yet compete. The exponential increase in information, knowledge, literacy, productivity and efficiency we see in social progress neither results from, nor translates into, an increased capacity for intelligence, imagination and consciousness. It results from the rising power to digitize, process and distribute information rapidly and globally.

The quality of multi-dimensional intelligence itself, though, is *not* growing. The profundity of thinking by the Ancient Greek philosophers, Aristotle and Plato, has probably never been surpassed. No one has had an impact on values, beliefs and consciousness comparable to the early founders of the major human religions. No writer since Shakespeare has come anywhere near his level of literary and imaginative genius. At the same time, no contemporary musician is claiming to have put Bach, Mozart and Beethoven to shame. It's unlikely we'll see the likes of scientists on the level of Isaac Newton and Albert Einstein any time soon.

I cannot see how our current evolution as a species is heading towards the creation of any super-intelligence, either in the form of self-aware, perceptually alive robots which could outwit and outmanoeuvre their human masters, or of some disembodied form of super-intelligence which could be controlled and centralized into a form of single, self-aware consciousness.

But a new species fusing technology and humanity to become super-beings and elites of a global technocracy – this *is* conceivable, possible and doable. I cannot support, though, the concept of any post-human future from a philosophical, ethical and social perspective.

Kurweil, though, argues that by mid-century, humanity will be indistinguishable from our technology, presupposing we will fuse completely with our technology, and that by the end of this century "the nonbiological portion of our intelligence will be trillions of trillions of times more powerful than unaided intelligence".[3] Within decades, he claims, information-based technologies will encompass all human knowledge and proficiency, including the emotional and moral intelligence of the human brain.[4] At this stage, machine intelligence will have drawn equal with human intelligence, but will have the added strengths of machine intelligence.

As the exponential growth of technology continues, however, machine intelligence will pull ahead of human intelligence until "most of the intelligence of our civilization will ultimately be non-biological".[5] Having conquered civilization, machine intelligence, according to Kurzweil, will then advance to conquer the universe itself (and, you guess it, any multiverses out there): "Ultimately, the entire universe will become saturated with our intelligence…we will determine our own fate rather than have it determined by the current 'dumb', simple, machinelike forces that rule celestial mechanics."[6] What he forgets in this brand of grandiose technological mysticism is that he has already airbrushed "we", that is, human beings (at least as we know them), out of the future control of civilization. But wait, did you just digest that last quoted statement? So the mathematics of celestial mechanics, which has kept the planets, solar system and galaxies in place for *billons* of years, and which enabled NASA to land Apollo 11 on the moon and bring its astronauts safely back to earth in 1969, are too "dumb" for such a vision of ultra-intelligence? I'm afraid this kind of speculation simply goes beyond the bounds of rational discourse.

What we *actually* have is not greater-than-human super-intelligence but the global brain of internet and the worldwide web coupled with powerful logical and methodological tools provided by science, the great problem-solver, with its immense treasury of knowledge.

We also have the foundations of civilization in place. The possibilities for social progress going forward through the rational application of scientific, logical and mathematical methods to improve the human condition are limitless.

---

[3] Kurzweil (2005): 9.
[4] Kurzweil (2005): 8.
[5] Kurzweil (2005): 30.
[6] Kurzweil (2005): 29.

Why and how would we lose control of this progress when science has, by and large, mapped out the laws that govern how things work? We can control and govern the technologies we create, while they themselves are subject to laws and limits of nature, just as we are. Not having human or organic form, machines seem condemned to stay well below self-aware, interpretative, creative, holistic human consciousness.

If rationality remains the prevailing mode of public discourse, human intelligence will bring about continuous social progression through the application of science without having to hand over governance to machines, robots or computer systems.

On that concluding note, I leave you with the words of the first man to fly in space, Soviet cosmonaut Yuri Gagarin, a rather remarkable individual by any standards. In his autobiography *Road to the Stars*, he describes how his generation of young Russian scientists, engineers and pilots had started to ask the question whether machines would one day replace humans. This is what he concluded: "By that time I had heard something about cybernetics and had heard people say that electronic machines would replace the human brain. There are many reasons for not agreeing with this idea, not the least of them the fact that man would always have to make decisions no matter how perfect the machine and that in critical cases man is more versatile….the human brain is nature's most perfect work, there is nothing to replace it and never will be."[7]

Acknowledgments

Gagarin, Y. (2002 reprint). *Road to the Stars*. Honolulu: University Press of the Pacific.

Good, I.1965. "Speculations Concerning the First Ultra-intelligent Machine". <u>Advances in Computers</u>(Volume 6).

Kaku, M. 2014. *The Future of the Mind*. New York: Doubleday.

Kurzweil, R. 2005. *The Singularity is Near – when humans transcend biology*. London:Duckworth.

More, M. & Vita-More, N, ed. 2013. *The Transhumanist Reader*. Chichester: Wiley-Blackwell.

Prince, R. 2014. The Telegraph. June 25, 2014 http://news.nationalpost.com/2014/06/25/science-fiction-come-true-paralyzed-man-moves-arm-with-his-thoughts-using-microchip-in-his-brain/

---

[7] Gagarin (2002):91.