Computers will not acquire general intelligence, but may still rule the world

RAGNAR FJELLAND
Professor Emeritus, University of Bergen

Jobst Langrebe's and Barry Smith's book *Why Machines Will Never Rule the World* argues that artificial general intelligence (AGI) will never be realized. Drawing on theories of complexity they argue that it is not only technically, but mathematically impossible to realize AGI. The book is the result of cooperation between a philosopher and a mathematician. In addition to a thorough treatment of mathematical modelling of complex systems the book addresses many fundamental philosophical questions. The authors show that philosophy is still relevant for questions of information technology in general and artificial intelligence in particular.

This paper endorses Landgrebe's and Smith's arguments that artificial general intelligence cannot be realized, but not their conclusion that machines will never rule the world. It is not only a question of what technology can do. An equally important question is what technology does to *us*. Machines may not take over the world in a literal sense, but they may have many negative effects. Some of the most serious can be placed under the category of the "degeneration effect".

INTRODUCTION

The play R.U.R. was written by the Czech writer Karel Capek in 1920. R.U.R. stands for "Rossum's Universal Robots". Capek was actually the first who used the word "robot" to denote a technical device, and he raised hundred years ago many of the issues that are central to today's debate on artificial intelligence.

Rossum is the name of the person in the play who invented the robots. The robots replace humans in production, but in the construction of the robots, everything superfluous has been left out. People feel happiness, they play the violin, like to take a walk and many other useless things. The robots are unable to do any of this. They are constructed according to the principle that the simplest is the best. In the play it is pointed out that the best worker is the cheapest worker, the one with the fewest possible needs. Therefore, the robots have no use for intelligence either. You can read a twenty-volume encyclopedia to them, and they can repeat it word for word, but they don't come up with anything new themselves: "They would make very good university lecturers", the text says (Capek 2015, p. 18).

Rossum's robots become a success, and form the basis for an industrial adventure. The factory, located on an island, produces thousands of robots that are sold all over

the world. The robots are produced by other robots, and only the top management in the factory consists of humans. However, even the best plan can go wrong. Although most robots are only equipped with the minimum features necessary to perform the tasks, it has been necessary to produce a few robots that are more intelligent than the others. They organize a rebellion and decide to exterminate all humans. As one of the leaders says: "To become like humans, it is necessary to kill and to dominate. Read the history books" (Capek 2015, p. 84)

The factory is surrounded by robots, but the managers, who are trapped there, hope that they will negotiate an agreement that the robots can take over the factory in exchange for them being allowed to escape in a boat. One of the managers regrets that he hasn't enjoyed life earlier, and exclaims: "Have fun. Beauty. Hell, there are so many beautiful things around us! The world was beautiful and we... we here.. tell me, what did we even appreciate?" (Capek 2015, p. 63). They then continue to fantasize about having a small farm, where they can live a peaceful country life with fresh air. But they never get that chance, because everyone is killed.

Capec was not the first to warn against the possibility that technology takes command. Mary Wollstonecraft Shelley had already written *Frankenstein; or, the Modern Prometheus* in 1818, and Lewis Mumford had written *The Pentagon of Power* in 1964. In 2014 the physicists Stephen Hawking, Max Tegmark and Frank Wilczek, and the computer scientist Stuart Russell, published an article in which they warned against what the development of artificial intelligence may lead to. In the article they said, among other things:

> Success in creating AI would be the biggest event in human history. Unfortunately, it might also be the last, unless we learn how to avoid the risks (Hawking 2014).

What Hawking and his co-authors warned against was the the possibility of creating artificial general intelligence (AGI).

In 2023 Future of Life Institute published an open letter calling for "all AI labs to immediately pause for at least 6 months the training of AI systems more powerful than GPT4". The letter was originally signed by prominent AI researchers (among others Stuart Russell), entrepreneurs (among others Elon Musk) and scholars (among others Yuval Noah Harari).[1] The open letter was prompted by the success of OpenAI's systems ChatGPT. In the letter it is not asserted that AGI has been realized, but it is argued that "... [c]ontemporary AI systems are now becoming human-competitive at general tasks". If one looks at the references, it looks as if the letter assumes that AGI may be realized in the near future. OpenAI shares the idea that GPT-4 is about to realize AGI. On their home page the headline says: "Our mission is to ensure that artificial general intelligence—AI systems that are generally smarter than humans—benefits all of humanity."[2]

Jobst Landgrebe's and Barry Smith's book *Why Machines Will Never Rule the World* (2023) directly addresses this question. They argue that there is no cause to worry about computers taking over the world. Their basic argument against the possibility of artificial general intelligence is simple: To construct artificial intelligence that simulates a system, we have to construct a mathematical model. It is impossible to construct complete mathematical models of complex systems. Most natural systems, including the human brain, are complex. Therefore, artificial general intelligence (AGI) is not only technically, but mathematically, impossible (Landgrebe and Smith 2023, p. 9).

Landgrebe is mathematician and artificial intelligence entrepreneur and Smith is philosopher. The book shows that philosophy is still relevant for questions of information technology in general and artificial intelligence (AI) in particular. It addresses many fundamental philosophical questions, like realism, reductionism, consciousness, language and ethics. The treatment of these philosophical topics is highly interesting for their own sake, independently of their significance for AI research. The same applies to the book's treatment of mathematical modeling as well. The strength of the book is that it does not just address the questions at a general and abstract level, but pursues questions into the technical details of AI research. This is the real mark of scholarship.

The book also brings a historical perspective into a field where the historical dimension has been missing. For example, they show that Aristotle is relevant for modern biology. It also brings information that turned out to be new and interesting to me. One example is that Karl Bühler, the teacher of both Karl Popper and Konrad Lorenz, emphasized the importance of language as an integrated part of human behavior, in particular that speech is an action in itself (Landgrebe and Smith 2023, p. 65). Another interesting piece of information is that Ilya Prigogine, who was a pioneer in the development of irreversible thermodynamics (which earned him a Nobel prize in chemistry) and the theory of complexity, was originally inspired by the philosopher Henri Bergson (Landgrebe and Smith 2023, p. 125).

WHAT KIND OF PHILOSOPHY?

How can philosophy shed light on the debate about AGI? After all, it looks as if the arguments in support of AGI are based on science. Probably the dominating view among researchers working on AI is that philosophy has nothing to contribute. Science and technology have made progress because they liberated themselves from philosophy during the scientific revolution. This view has been accepted by many philosophers as well. For example, Newton's contemporary John Locke declared himself an "under-laborer" in comparison to Newton (Locke 1971, p. 58), (Burtt 1972, p. 18). He was followed by other philosophers, not least the logical positivists in the last century. For example, the prominent logical positivist Hans Reichenbach echoed Locke when he stated that the only thing "the philosopher can do is to analyze the results of science, to construe their meaning and stake out their validity" (Reichenbach 1949, p. 312).

However, the arguments supporting the possibility of AGI are more based on metaphysics than science. Modern science did not completely liberate itself from metaphysics. This is obvious if we turn to one of the leading figures of the scientific revolution, Galileo Galilei. He was not only a scientist, but a philosopher of science as well. Although he no doubt regarded them as two sides of the same coin, in hindsight we may separate them. Some of his basic philosophical assumptions were published as early as 1623, in *The Assayer*, where he compared the universe to a grand book that is written in the language of mathematics, and "its characters are triangles, circles, and other geometric figures without which it is humanly impossible to understand a single word of it" (Galilei 1957a, pp. 237-8). He explains that when he conceives of any material or corporeal substance, he thinks of it as bounded and having a specific shape, being large or small, being located in a specific place at a specific time, being at rest or in motion, as touching or not touching another body and being one or many. All these properties can be described mathematically, and they are real. What cannot be described mathematically, is not real. This applies to tastes, odors and colors, which

> ... reside only in the consciousness. Hence if the living creature were removed, all these qualities would be wiped away and annihilated (Galilei 1957, p. 274).

According to Galileo's younger contemporary Descartes the things that exist in the objective world are characterized by extension, in Latin they are res extensa. This is nothing but metaphysics.

Another basic idea in Galileo's metaphysics is that there is no real difference between the natural and the artificial. According to Descartes "...the laws of mechanics ...are the same as those of nature..." (Descartes 1973, p. 115). In other words, the world is a machine. At the time of Galileo and Descartes the paradigmatic machine was the clock. Today it is the computer.

As the philosopher Edmund Husserl pointed out in his late work *The Crisis of European Science and Transcendental Phenomenology*, Galileo was "... at once a discoverer and a concealing genius" (Husserl 1970, p. 52). On the one hand he established the ideal of mathematical science, but on the other hand he substituted the "mathematically substructed world of idealities" for our everyday lifeworld. This was Galileo' metaphysics, and it was passed on to Galileo's successors, "the physicists of all the succeeding centuries" (Husserl 1970, p. 49). The consequence was a fundamental *misunderstanding* of modern science. In other words,

Husserl endorsed Galileo's physics, but regarded his metaphysics as a misunderstanding. Galileo's misunderstanding was rooted in his Platonism. Husserl called it "objectivism", "physicalism" and "naturalism".

Husserl influenced many "continental" philosophers, not least Martin Heidegger and Maurice Merleau-Ponty. The philosopher Hubert Dreyfus was in his turn influenced by them. In 1972 he launched an attack on mainstream assumptions in AI research in his book *What Computers Can't Do* (Dreyfus 1972). He argued that mainstream AI research was based on untenable metaphysical assumptions. The bottom line of his argument was that computers, who have no body, no childhood and no cultural practice, cannot acquire general, or human-like, intelligence. However, at the time when Dreyfus wrote his book artificial intelligence was what the philosopher John Haugland called "Good Old-Fashioned AI" (GOFAI). Stuart Russell and Peter Norvig have therefore argued that a more adequate title of Hubert Dreyfus' book would have been "*What First-Order Logical rule-Based Systems Without Learning Can't Do*" (Russel and Norvig 2016, p. 1024). Their point is that AI research has made tremendous progress since Dreufus wrote his book, in particular based on artificial neural networks, and that this development has rendered most of his arguments invalid or inapplicable.

Landgrebe's and Smith's book fits into the tradition from Husserl, Heidegger, Merleau-Ponty and Dreyfus. Although their approach is different from Dreyfus', and they only mention him in the passing, they show that Dreyfus was basically right.

LIMITS: COMPLEXITY

In the introduction I referred to scientists and entrepreneurs who warned against the dangers of AI getting out of control. Future of Life Institute describes its own mission with the following words:

> Steering transformative technology towards benefiting life and away from extreme large-scale risks. We believe that the way powerful technology is developed and used will be the most important factor in determining the prospects for the future of life. This is why we have made it our mission to ensure that technology continues to improve those prospects (https://futureoflife.org/our-mission/ (accessed 16.6.2023)).

However, probably none of the signatories of the the institute's open letter are against AI. Many of the researchers and entrepreneurs in the field stress the benefits more than the risks, and some have grand visions of the future of AI. Max Tegmark, president of Future of Life Institute refers to a late-night discussion he had with Larry Page, co-founder of Google. According to Tegmark, Page gave "a passionate defense of the position I like to think of as *digital utopianism*." This position entails that digital life is the natural and desirable next step in the cosmic evolution, and if we let digital minds be free, the outcome is almost certain to be good (Tegmark 2017, p. 32).

Jaron Lanier has also been an insider in the Silicon Valley community. Among other things, he is one of the founders of the field of virtual reality. In his book *Who Owns the Future*, he claims that the dominant view in leading information technology circles in Silicon Valley is technological determinism: The basic assumption is that at some point in the not-too-distant future (we are talking about a few decades), the Internet will develop into an artificial superintelligence that will be much more intelligent than any human. It will take over the world before the majority even notice it. Lanier says that such thoughts can seem crazy when presented in this way, but nevertheless they are widespread in the circles he knows, and argues that "these are guiding principles, not just amusements, for many of the most influential technologists." He goes so far as to claim that we have a new religion, which is expressed in a technological culture (Lanier 2013, p. 186).

Needless to say, Landgrebe and Smith are opposed to such ideas. As already mentioned, the core of their book is that complex systems cannot in general be modeled mathematically. They show that many of

the arguments in favor of AGI are not based on science, but rather science fiction. I will go through two of their many examples: the emulation of human language and the emulation of the human brain.

COMPUTERS WILL NEVER BE ABLE TO EMULATE HUMAN LANGUAGE

Aristotle defined man as the rational animal, and the distinguishing mark of rationality was language. In the article "Computing Machinery and Intelligence" Alan Turing asked the question: How can we actually determine if computers have acquired general intelligence? (Turing 1950) He starts by saying that the question he is trying to answer is: "Can machines think?", but instead of going into the question of what intelligence is, he sets up a kind of game. In the game a questioner can communicate with a computer and a human being. He has to communicate through a keyboard, so he does not know who is the computer and who is the human. The point is that the machine pretends to be human, and it is the job of the questioner to decide which of the two is the computer and who is the human. If the questioner is unable to distinguish, we can say that the computer is intelligent. Turing called this the "imitation game", but it is later known as the "Turing test". If the computer passes the test, it has, according to Turing, acquired general intelligence. Therefore, the Turing test is about human language.

Joseph Weizenbaum, at that time professor of informatics at MIT, created the first natural language processing computer program. He named it ELIZA, after Eliza Doolittle, the main figure in George Bernard Shaw's *Pygmalion*. The program could be taught to "speak" increasingly well, but like Eliza Doolittle, it was not clear whether it became smarter. (Weizenbaum 1976, 188n) The program enabled the user to carry out a "conversation" with a computer using everyday language. Weizenbaum himself was fully aware of the fact that the computer did not understand anything. Therefore, he was surprised when he learned that the program created the illusion among people who used it that the computer really understood, in particular among people who did not know anything about computers (Weizenbaum 1976, p. 189).

Today there is a lot of hype related to the presentation of OpenAI's ChatGPT and GPT-4. Although the performance is impressive, it is interesting to see that the weaknesses are the same as Hubert Dreyfus pointed out fifty years ago. This is well illustrated in Landgrebe's and Smith's example of the problem of computing the appropriate length of a pause in a conversation. A pause length depends on context (for example, is it a memorial dinner, a cocktail party, or an argument among Parisian intellectuals), it may depend on the emotional loading of the situation, on knowledge of the other person's social standing, or something as simple as the other person consults his phone. Therefore, the appropriate length of a pause depends on an understanding of the whole situation (Landgrebe and Smith 2023, p. 243).

Language is communication, and fundamentally the primary communication situation is face-to-face. Landgrebe and Smith show that language at all levels, from the primary situation of face-to-face communication, involves complex systems. In general they cannot be modeled mathematically. Some sub-systems, or some aspects of complex systems, can be modeled. And these can be emulated by AI. However, there will never be complete AI language systems, because even simple everyday conversations presuppose a multitude of complex systems at different levels.

One might argue that this is too much to expect from a chatbot, because we cannot communicate with a chatbot "face-to-face". However, the defect is deeper. Because a computer is not in the world, it cannot in principle distinguish between an object and the representation of an object, or between reality and appearance (Smith 2019, p. 81ff). Chatbots like ChatGPT will therefore always remain unreliable.

COMPUTERS WILL NEVER BE ABLE TO EMULATE THE HUMAN BRAIN

What looks as a straightforward way to creating AGI would be to emulate the human brain. The most prominent advocate of this view is the philosopher David Chalmers. The core of his argument is that the brain is a machine, and, therefore, we will be able to emulate it "before long". Landgrebe and Smith also quote

Anders Sandberg who has said that according to current neuroscientific and technological knowledge there are no fundamental obstacle to "whole brain emulation".

If the brain had been a machine, this argument would have made sense. In fact, the most successful application of mathematical models is in the design of machines. However, Chalmers and Sandberg fail to see the essential difference between a machine and an organism, including the human brain. Landgrebe and Smith point out that we have descriptive models of some biological subsystems, as well as explanatory and predictive models of some very small subsystems. But we cannot make a mathematical model of even the most primitive living organism, like *archaeum*, because it is a complex system involving the dynamic interaction of more than 100000 biomolecules (Landgrebe and Smith 2023, p. 198). Their conclusion is that there is no scientific support for the project of emulating the human brain.

THE SINGULARITY

The most common argument in favor of AGI is based on the development of supercomputers and even hypercomputers. For example, Raymond Kurzweil, technical director of Google, argue that most long-range forecasts underestimate the speed of future technical development because they are based on "the intuitive linear" view of history rather than the "historical exponential" view. His predictions about future technical development are based on the "law of accelerating returns". In other words, the development is exponential.

Kurzweil's basic concept is the Singularity. The word singularity, which was first introduced in this context by Victor Vinge, is taken over from mathematics and physics. In physics, it denotes a state where the laws of physics break down, such as at the beginning of the universe (the "big bang") and in a black hole. In Kurzweil's sense, it is the point at which computers have the same intelligence as humans. He estimates this time to be approximately the year 2045. In Kurzweil's words:

> The Singularity will enable us to transcend the limitations of our bodies and our minds. We will gain control of our own destinies. Our mortality will be in our own hands. We will be able to live as long as we want… (Kurzweil 2013, p. 9).

After the Singularity, the intelligence of computers will continue growing exponentially, and they will be millions of times more intelligent than us.

But if computers are millions of times as intelligent as humans, why will they want to be our servants? I started with Capec's *R.U.R.* Before the robot revolt the most intelligent robot, Radius, has a conversation with a member of the management, Helena. She reminds him that he has been constructed with a brain that is twice as big as a human brain:

> Helena: That's why I had you put in the library, so that you could read up on everything. Oh, Radius, I wanted you to show the world that robots are as good as we are.
>
> Radius: I wish to have no master.
>
> Helena: Nobody would give you orders. You'd be just like us.
>
> Radius: I wish to be the master of others (Capek 2015, p. 44).

This is the Terminator problem. According to Landgrebe and Smith there is no reason to fear this scenario. In the first place, robots will not acquire superintelligence. Second, they will not have their own will (Landgrebe and Smith, p. 277).

COMPUTERS MAY STILL RULE THE WORLD

When I say that computers may still rule the world, it should not be taken literally. Machines will never rule the world in the way Capec described in *R.U.R.*

According to Landgrebe and Smith machines will never rule the world because they cannot acquire general intelligence. But is general intelligence required to rule the world? If we look at politicians of the last decade we may ask the question how much intelligence is needed to rule the world. Can we exclude the possibility that idiots may rule the world? Unfortunately, I think the answer is no. However, I shall not pursue that question further.

I have previously emphasized that the roots of information technology and AI can be traced back to the scientific revolution of the seventeenth century, and pointed out that one of the distinguishing marks of the revolution was the disappearance of the distinction between the natural and the artificial. Although the idea that the natural and the artificial are identical is pure metaphysics, it has prevailed until this day (in spite of the criticism by for example, Husserl, Heidegger, Merleau-Ponty and other philosophers). As a consequence, many technologists do not think that they simply *invent* things. They rather *discover*. Another consequence is a kind of technological determinism, in the sense that the inventions that are made, could in principle not have been different.

I have previously referred to what Tegmark called "digital utopianism". The most interesting formulation of this ideology can be found in Kevin Kelly's book *What Technology Wants* (Kelly 2010). Kelly is one of the founders of the leading magazine *Wired*, and was its editor for the first seven years. The journal was started in 1993, with the aim of dealing with the importance of the new technology for economics, politics and culture. He has summarized his view in a number of theses published in the blog *Cooltools* that he started. Here are some of his theses:

- The progression of technologies is inevitable.
- Because technologies are inevitable we can prepare to optimize their benefits.
- Technology is not neutral but serves as an overwhelming positive force in human culture.
- We have a moral obligation to increase technology because it increases opportunities.
- The origins of technology lie in the Big Bang.
- Technology preceeded humans and will continue beyond us. (https://kk.org/cooltools/what-technology/(accessed 21.6.2023))

Needless to say, this is metaphysics. However, as I have indicated, it can be fitted into a tradition that goes back to the scientific revolution. We may argue that this kind of technological determinism is untenable, but the problem is that the view is shared by many of the leading persons in Silicon Valley. Therefore, I think I am justified in arguing that computers may rule the world even if they do not acquire general intelligence.

Landgrebe and Smith have a subheading: "How AI will change the world." The paragraph is a little more than half a page, and represents the end of the book. They argue that the main challenge in the future is to find new occupation to those whose jobs have been mechanized. However, they do not regard this as a serious threat, because they "... are confident that, as in the past, new occupations will evolve in ways that no one (and no algorithm) could have predicted" (Landgrebe and Smith 2023, p. 301). When we take into consideration that the subtitle of the book is "Artificial Intelligence without Fear", it is remarkable that they do not see more possible negative effects of AI.

The Future of Life Institute's open letter rightly points to some risks that should be taken seriously:

> Contemporary AI systems are now becoming human-competitive at general tasks, and we must ask ourselves: Should we let machines flood our information channels with propaganda and untruth? Should we automate away all the jobs, including the fulfilling ones? Should we develop nonhuman minds that might eventually outnumber, outsmart, obsolete and replace us? Should we risk loss of control of our civilization? Such decisions must not be delegated to unelected tech leaders.

Some of the things the letter warns against, for example machines flooding our information channels with propaganda and untruth, or automating fullfilling jobs, are real threats to society. They should be taken seriously, even if we disagree with the allegation that AI systems are about to become "human-competitive at general tasks". The list could also be made longer, by adding, for example, surveillance. We know that AI is used extensively for surveillance in China. However, it does not only apply to totalitarian societies, but applies to allegedly democratic societies as well. This is well documented in Shoshana Zuboff's brilliant book *The Age of Surveillance Capitalism* (Zuboff 2019).

Let me return to R.U.R. Capec saw something important: That technology does something to us. When the managers of the robot factory were surrounded by robots, they recognized that they had missed something. They had been so busy producing and selling robots that they had forgotten what is important in life. Although Capec pointed to an important problem, he did not see the real problem: How technology changes us. One thing is what computers can do and can't do. Another thing is what technology in general, and information technology and AI in particular, does to *us*, as individuals and as society. A visible indication of the impact is the fact that the cell phone has changed the behavior and interactions of a large part of the world's population. Millions, or rather billions, of people walk around with a cell phone in one hand, and a large part of them look at the display as they walk. We also know that a large parts of them are addicted. Adam Alter starts his book *Irresistible* quoting Steve Jobs when he presented the iPad in 2010 and emphasized its extraordinary properties. He used words like "extraordinary", "incredible experience" and "phenomenal". Alter then adds the dry remark: "But he refused to let his kids use the device" (Alter 2017, pp. 1-2).

I will restrict myself to pointing to the way AI changes our acquisition of skills and knowledge, and even our conception of knowledge.

THE DEGENERATION EFFECT

The "degeneration effect" denotes the phenomenon that skills that are not practiced, tend to degenerate. Jonathan Carr gives an illustrating example in *The Glass Cage* (2015). The example is the flight from Newark to Buffalo on February 12, 1989. The plane took off on manual control, and when airborne, the pilots engaged the autopilot. When, after an hour, it approached the airport in Buffalo, there was a signal that the plane was losing lift, and risked going into an aerodynamic stall. The autopilot automatically disengaged, as it was programmed to do, and the captain took over. He reacted quickly, but did the opposite of what he should have done. He pulled the lever towards him, so that the plane went up and slowed down. Instead, he should have pushed the lever forward, so that the plane turned downwards and increased its speed. The plane crashed into a house, killing all forty-nine on board.

Carr gives many more examples of how traditional skills may degenerate. The core of the problem is that people get used to blindly trusting technology, so that they switch off their traditional skills and even ignore clear danger signals. Another example is the passenger ship "Royal Majesty", which in the spring of 1995 was on its way from Bermuda to Boston. It was equipped with the most modern automatic navigation equipment that used GPS to determine the course. But after an hour's sailing from Bermuda, the antenna came out of position. The GPS was still giving signals, but they were incorrect so that the ship gradually drifted off course. The ship sailed on for thirty hours without the crew noticing that it was off course. At one point, the officer on duty did not observe an important buoy that the boat was about to pass, but he relied so much on the technology that he failed to report it. It ended with the ship going on a sandbar. No one was injured, but the company suffered heavy losses.

READING

Plato already knew that all technological progress comes at a price. In his dialogue *Phaedros* he tells the myth about the Egyptian god Teuth, who among other things had invented the alphabet. Teuth describes the advantages of writing to the Egyptian king Thamus, and claims that it will improve the memory of the people of Egypt. Thamus disagrees, and argues that the effect of the invention will be the opposite of what Teuth claims: Relying too much on written language will impair memory. Therefore, Thamus claims, Teuth had invented a technology of forgetfulness. (Plato 1973, p. 275).

A more recent example is given in David Pogue's monthly column in *Scientific American* for August 2013. He recounts that when his father grew up, his father (Pogue's grandfather) offered him 25 cents if he could memorize the complete list of former US presidents. His father in his turn offered Pogue a generation later 5 dollars if he could do the same. The increase in reward was justified by inflation and the increase in the number of presidents. Pogue himself offered his son 10 dollars if he could do the same. However, his son was baffled, because he could not understand why he should memorize former presidents, and argued that "everybody has a smartphone". Pogue himself agrees:

> In other words, having a computer in your pocket is the norm. Google is always one tap away. So there's little sense, as far as my son is concerned, in memorizing anything: presidents, the periodic table of the elements, the state capitals or the multiplication tables above 10 (Pogue 2013, p. 25).

Pogue has an important point: It is a waste of time and cognitive capacity to memorize the entire list of American presidents, the entire periodic table and state capitals. In this regard, an impaired memory is a small price to pay for the advantages of written language, not to say information technology and the Internet.

However, the price of progress that Plato pointed to, is real, and applies to all technology. Information technology and the Internet are no exception, and Nicholas Carr in his book *The Shallows* (Carr 2010) has pointed to some serious negative consequences. At the beginning of the book he quotes a philosophy major student who says that it makes no sense to go through a book from cover to cover "...as I can get all the information I need faster through the Web" (Carr 2010, p. 9). Put in oversimplified form Carr's thesis is that when Google has scanned the last book, nobody reads books anymore. The "deep" knowledge that can only be conveyed by a book, will disappear, because the Internet fundamentally changes the way we read: We tend to search, and retrieve fragmented knowledge. If we know what we are looking for, this is sometimes very useful. But if we lack the background knowledge, or the context, it is dangerous. The danger is that the reader becomes more ignorant, because he thinks he knows something, although that knowledge is so fragmented and shallow that it can hardly be called knowledge.

WRITING

Reading and writing are two sides of the same coin, and Pogue's son's argument that it does not make sense to memorize the entire list of former presidents because everybody has a smartphone in his or her pocket may be applied to writing as well. Why write tedious texts when ChatGPT or another chatbot is at hand, and can do it much faster and better?

ChatGPT and similar systems will create new problems with Ghost-writing and plagiarism. This is not the main problem, as I suppose that in the near future there will be available programs that can reveal machine-produced texts. But if we disregard possible abuse, and assume that ChatGPT is used as a tool, what are probable consequences?

To see the problem it is useful to compare to mathematics. If we just learn some procedures or mathematical expressions, we have not learned mathematics. To learn mathematics we have work through exercises and proofs. It is an active process. Of course, this comes in degrees. For example, we may apply the for-

mula for a normal distribution without being able to derive it. However, if we are only able to apply formulas and follow procedures that we don't understand, there is always the danger that we may make serious errors or abuse the little knowledge that we have.

The same applies, in various degrees, to other fields of knowledge. Knowledge acquisition is an active process. Isolated facts and ideas do not represent real knowledge, because they lack coherence. For something to become real knowledge, we have to work through the facts and ideas that we acquire. To create a text "from scratch" is an important part of this process. It is worth keeping in mind that a well-known indication of lack of real knowledge or understanding is the inability to give examples or to account for a problem in one's own words.

When we write, we normally use various sources of information, like articles, books, encyclopedias or the Internet. However, the Internet makes a difference. It is at "our fingertips", and it is, therefore, easy to copy a sentence from one source, and another sentence from another source. Normally this will not qualify as plagiarism. However, it comes at a price, because the text will be broken up and may often lack coherence. My own advice to students has always been that they should try to write a text "in their own words". When this advice is followed, the effect is sometimes a dramatic improvement in the quality of the text.

If we lose the ability to read longer texts and write independently, then we will have a problem, and the situation will even be more serious if we are not aware of the problem. Having search engines and chatbots permanently available may give us the illusion that we are omnipotent and master the world.

The economist Eli P Cox III has used the expression "shadow curriculum" (inspired by Émile Durkheim's expression "hidden curriculum") which represents the values and beliefs that are embedded in an educational setting (Cox III 2017, p. 65). Social media no doubt have a shadow curriculum, and it is brilliantly described in Johann Hari's book *Stolen Focus*. For example, the shadow curriculum of Twitter is that "[t]he world can and should be understood in short, simple statements of 280 characters." Twitter may be regarded as extreme, but nevertheless, a former US president for some time used it as his main source of communication. As Hari himself remarks, the problem is that the world is complex (Hari 2023, pp. 79-80).

And, of course, this is Landgrebe's and Smith's main point as well.

CONCLUSION

I have pointed out that Landgrebe's and Smith's book *Why Machines will Never Rule the World* is important and brilliant. I have also criticized what I regard as an important omission. However, this defect can easily be remedied by changing the title, and removing the subtitle and the last page of the book.

NOTES

1 As of 15.6.2013 the letter has 31810 signatures. https://futureoflife.org/open-letter/pause-giant-ai-experiments/

2 https://openai.com/blog/planning-for-agi-and-beyond#SamAltman

REFERENCES

Alter, Adam. 2017. *Irresistible. The Rise of Addictive Technology and the Business of Keeping Us Hooked.* New York: Penguin Press.

Burtt, Edwin Arthur. 1972. *The Metaphysical Foundations of Modern Physical Science* (1924). London: Routledge Kegan Paul.

Capek, Karel. 2015. *R.U.R* (*Rossum's Universal Robots,* 1920). Unknown.

Carr, Nicholas. 2010. *The Shallows. How the internet is changing the way we think, read and remember.* London: Atlantic Books.

Cox III, Eli P. 2017. *Seeking Adam Smith. Finding the Shadow Curriculum of Business.* Singapore: World Scientific Publishing.

Descartes, René. 1973/1637. Discourse on Method. In: *The Philosophical Works of Descartes*, Tr. S. E. Sutcliffe. Vol. I. Cambridge: Cambridge at the University Press.
Dreyfus, Hubert L. 1972. *What Computers Can't Do*. New York: Harper and Row.
Galilei, Galileo. 1957/1623. The Assayer (Il Saggiatore). In: *Discoveries and Opinions of Galileo*, ed. Stillman Drake. New York: Doubleday & Company.
Hari, Johann. 2023. *Stolen Focus. Why You Can't Pay Attention*. London: Bloomsbury Publishing.
Hawking, S., M. Tegmark, F. Wilczek and S. Russell. 2014. Transcendence Looks at the Implications of Artificial Intelligence — but Are We Taking Ai Seriously Enough? *Independent*. 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.
Husserl, Edmund. 1970. *The Crisis of European Sciences and Transcendental Phenomenology*. Evanston: Northwestern University Press.
Kelly, Kevin. 2010. *What Technology Wants*. Harmondsworth: Penguin Books.
Kurzweil, Ray. 2013. *The Singularity Is Near*. London: Duckworth.
Landgrebe, Jobst, and Barry Smith. 2023. *Why Machines Will Never Rule the World. Artificial Intelligence Without Fear*. New York, London: Routledge.
Lanier, Jaron. 2013. *Who Owns the Future?* London: Allen Lane.
Locke, John. 1971/1690. *An Essay Concerning Human Understanding*. Ed. A. D. Woozley. Glasgow: Fontana/Collins.
Plato. 1973. *Phaedrus and Letters Vii and Viii*. Harmondsworth: Penguin Books.
Pogue, David. 2013. The Last Thing You'll Memorize. *Scientific American*, 25.
Reichenbach, Hans. 1949. The Philosophical Significance of the Theory of Relativity. In: *Albert Einstein: Philosopher-Scientist*, ed. Paul Arthur Schilpp, 1:287-312. La Salle: Open Court.
Russel, Stuart and Peter Norvig. 2016. *Artificial Intelligence. A Modern Approach. Third Edition*. 3rd ed. Prentice Hall Series in Artificial Intelligence. Harlow: Pearson Education Limited.
Smith, Brian Cantwell. 2019. *The Promise of Artificial Intelligence. Reckoning and Judgment*. Cambridge, MA; London: The MIT Press.
Tegmark, Max. 2017. *Life 3.0. Being Human in the Age of Artificial Intelligence*. New York: Alfred A. Knopf.
Turing, Alan. 1950. Computing Machinery and Intelligence. *Mind* LIX (236):433-60.
Weizenbaum, Joseph. 1976. *Computer Power and Human Reason*. San Francisco: Freeman and Company.
Zuboff, Shoshana. 2019. *The Age of Surveillance Capitalism*. London: Profile Books.