

The death of democracy through artificial intelligence

November 18, 2024 - [Andreas von Westphalen](#)

Translation with www.DeepL.com/Translator from:

<https://www.telepolis.de/features/Tod-der-Demokratie-durch-kuenstliche-Intelligenz-10031952.html>

Note on quotations: Some quotations are taken from German translations of English-language books, which have been translated from German into English using deepL and may therefore differ from the original.

Progress in AI raises hopes and fears. Technology is conquering our lives. But could AI become the gravedigger of democracy? (Part 2 and conclusion)

In the first part of this interview, Prof. Karl Hans Bläsius, expert in artificial intelligence at Trier University of Applied Sciences and operator of the websites "[Nuclear war by mistake](#)" and "[AI consequences](#)", discusses the dangers and control possibilities of AI and superintelligence.

► In your opinion, what problem do you think has been given too little consideration in discussions to date with regard to AI and how it will change our lives?

Karl Hans Bläsius: In the past, when people speculated about machines that could be superior to us in terms of intelligence, they mainly had in mind autonomous, intelligent robots that could plan and carry out actions.

However, the current successes of AI are all about chatbots that answer our questions on the internet. Systems such as ChatGPT are active on the internet in a certain sense "autonomously". They answer questions and solve problems without human intervention, and they can also program and be used for cyber attacks.

Once launched, such systems could search for potential targets and plan and execute attacks without further human intervention, meaning they can be considered autonomous cyber weapons. These systems could be misused by groups or individuals for cyber attacks on a country's critical infrastructure, for example.

Several generative AI systems similar to ChatGPT are already in use on the internet and more will be added.

Many companies and countries are currently working on generative AI systems. In addition to humans, bots can also pose questions and tasks to these systems and it is therefore to be expected that there will soon be interactions between these systems themselves.

This can lead to new dangers, especially if these systems have cyberattack capabilities.

When instructed by humans, bots or another generative AI system, a system like ChatGPT could carry out cyberattacks. Other generative AI systems, with which there are already interactions, could detect this and launch counterattacks. Without humans being involved, a chain reaction between these systems with ever stronger cyber attacks could develop in a short space of time.

It is known from the financial markets that unforeseen interaction processes can occur in high-frequency trading between different algorithms, which can lead to price crashes and financial losses within seconds, which is referred to as a "flash crash". Similar escalation spirals with cyber attacks would therefore also be conceivable through generative AI systems, which could be described as a "flash war" on the internet.

Even if there is no interaction between the current systems yet, this is likely to happen in the next few years or very soon. It is difficult to protect against these risks, because with a large number of such systems, it cannot be expected that the necessary security measures will be observed by all systems.

Furthermore, there could also be dangerous interactions between generative AI systems that are assigned to states that are currently on a confrontational course.

This could result in mutual recriminations and international conflicts. Such a "flash war" on the Internet could also be triggered by individual people or small groups, or such effects could arise by chance as a result of an unfavorable action.

The risks described here do not require the AI systems to become independent by having a mind of their own, pursuing their own goals or having a consciousness. None of this is necessary. The risks described result solely from the fact that suitable strategies and heuristics are required as a basis for good automatic problem solving, which enable solutions in huge search spaces.

To this end, it makes sense to evaluate previous actions and determine adjustments for the weights of possible operations. This alone can encourage escalating behavior and lead to chain reactions that could have serious consequences in a short time.

These risks can exist long before an AGI (Artificial General Intelligence) or superintelligence emerges. These are not necessary for such escalating behavior.

If these systems, possibly as a result of such interactions, dominate the flow of information on the Internet and thus paralyze human information flows, these systems would achieve information dominance that could affect many areas, including finance.

Due to today's dependence on the Internet, the consequences would be serious. Finance and trade could collapse, at least temporarily, and our social systems could become unstable.

Tight timeframe for decisions

► In your current article, you write:

Compared to other risks, such as climate change, it is completely incalculable whether, when and with what consequences a superintelligence could emerge. It is almost impossible to predict this. Such events are more likely to happen suddenly. Serious consequences could then occur within a few weeks or months, with no possibility of stopping them. The consequences could affect all or a large part of humanity.

These serious consequences could occur irreversibly within the next few years or decades. It may no longer be possible to wait for such events to occur, gain experience, obtain reliable knowledge

about the dangers and only then act to reduce the risks. Measures to reduce these risks would have to be taken beforehand.

Can you please explain this in more detail?

Karl Hans Bläsius: I see the greatest risk in the fact that the actions of generative AI systems could at some point lead to a breakdown in the exchange of information on the internet. If such events occur, this will happen relatively suddenly, possibly with serious consequences worldwide.

If finance and trade collapse, this could also lead to riots, civil wars or wars. To cause such effects, these systems do not need anything like consciousness or self-will, nor do they need to be close to a possible AGI.

If an AGI or superintelligence does eventually emerge, it may not be visible beforehand. This could be because a superintelligence, as described by Nick Bostrom, emerges in secret, and its power is not realized until it is present. It is not certain that a superintelligence will be realizable.

However, if one does emerge, it is completely unclear how this will take place, when it will be recognizable and to what extent, and when which effects will become visible. The emergence of a superintelligence will probably not be controlled by humans, but will occur suddenly and unexpectedly, possibly without the possibility of intervention and without the possibility of reversing anything.

Murray Shanahan, professor of cognitive robotics, writes in his book "The Technological Singularity" that any improvement in AI systems will be positive and that there will therefore be no reason to restrict such development.

At some point, however, a threshold could be reached at which these systems develop independently and humans no longer have the ability to intervene. To prevent this, measures would have to be taken beforehand. However, it is difficult to estimate when this would have to happen and how much time is left.

Abolition of the age of consent

► Immanuel Kant famously defined enlightenment as "man's emergence from his self-inflicted immaturity." Neuroscientist, doctor and psychotherapist Joachim Bauer warns in his latest book "Realitätsverlust" ("Loss of reality"): "Healthy people are helped to think until they can no longer think." Do you share his fears?

Karl Hans Bläsius: This fear is justified; Gabi Reinmann, Professor of Teaching and Learning at Universities at the University of Hamburg, also warns of such risks in a discussion paper published in October 2023 and refers to this as deskilling.

If more and more tasks are taken over by machines, the professional exchange from person to person will also be reduced and skills will be lost as a result. Technical aids are also increasingly being used at school and university to solve problems that used to be solved by the students themselves.

Systems such as ChatGPT can take over homework and make individual thinking superfluous, which will have a direct impact on our skills.

Technological progress has also created more and more dependencies on technical systems, especially electricity and now also the internet. Increasing capabilities of generative AI systems could lead to a dependency on these systems for knowledge and information. Humans will no longer need to know anything, as they can always ask AI systems for everything.

Artificial neural networks

► The name itself makes it clear that "artificial neural networks" refer to the "real" neural networks of the brain and are, as it were, modeled on them. In recent decades, however, research has shown that humans think and make moral decisions with mind and body, or more precisely, that thinking without a body is not possible. The neuroscientist Antonio Damasio writes pointedly in his book "The Strange Order of Things":

“The brain and the body are in the same boat and together they make the mind possible.”

Hence my question: to what extent is the name "artificial neural networks" accurate? And does it suggest that artificial intelligence is comparable to human intelligence, even if it is fundamentally a different kind of intelligence to that of humans?

Karl Hans Bläsius: I wouldn't disagree that thinking is not possible without a body. You could apply this analogously to the machine, because the software has to run on hardware in order to do anything. The term "artificial intelligence" can be understood to mean that the aim is to create systems that can solve problems for which humans need a certain degree of intelligence, although this can be achieved in completely different ways.

For me, this term did not imply that artificial intelligence has to be comparable to human intelligence. I also consider the term "artificial neural networks" to be appropriate, as this is an attempt to use knowledge about human thinking to implement solutions.

The current successes of AI are primarily based on artificial neural networks and deep learning. It is widely expected that these techniques will form the basis for the development of a superintelligence. But this is not necessarily the case.

Perhaps there are limits. Some AI scientists believe that different approaches need to be combined. Perhaps certain methods of symbolic AI will also be needed to achieve something like an AGI or superintelligence. We will only know exactly what will be needed once an AGI or superintelligence has been achieved.

Human understanding

► The question of what understanding actually is also crops up again and again in discussions about AI. Current research has found that interpersonal understanding always has a fundamental physical component, which is referred to as resonance or mirroring. Mirror neurons are fundamental to understanding between two people.

So to what extent does the understanding of AI remain fundamentally different from that of humans?

Karl Hans Bläsius: Systems such as ChatGPT have enormous capabilities in linguistic communication, although the question arises as to whether this can be described as understanding

natural language. Many people reject this in principle, with understanding being regarded as a binary property. This means that only the two possible values "applies" or "does not apply" are considered for understanding. However, this is not appropriate.

When we talk to children, a five-month-old child understands less than a five-year-old child. Understanding is a continuous process, whereby a certain level is always reached that lies somewhere between "applies" (100 percent) and "does not apply" (0 percent). Any value in between is possible. Understanding therefore has a vague character. It makes little sense to place systems such as ChatGPT at zero percent in terms of understanding.

"Understanding" will be fundamentally different for machines than for humans. The physical component and interpersonal understanding will be missing. The consequence could be that human needs are not adequately taken into account when a machine provides answers, recommendations or decisions, as the machine lacks the necessary empathy.

Horrendous energy consumption

► In the future, between ten and 20 percent of all electricity in the USA will probably be used solely for the development of AI. [Microsoft is considering building small nuclear power plants](#) to satisfy the hunger for energy for AI research.

Currently, data centers for training AI also consume [huge amounts of water](#) and, according to [a study](#), artificial intelligence could cause up to a thousand times more associated electronic waste in 2030 than was generated last year.

Is there a risk that AI, which is supposed to help make climate policy more efficient and reduce energy and resource consumption, will do the exact opposite and cause more harm than good?

Karl Hans Bläsius: Not all AI applications and learning processes are energy-intensive. Symbolic AI techniques do not always have to be based on large amounts of data. Learning procedures can also be realized with a few characteristic examples in some applications.

Whether a high energy requirement is necessary for normal AI applications depends on the respective application and the methods used.

The situation is different with generative AI. Systems such as ChatGPT are based on the analysis of huge amounts of data and require extremely high computing power. The energy requirement is correspondingly high. The further development of such systems therefore not only harbors fundamental risks for humanity, but will also place an enormous burden on the climate and/or increase the risks posed by nuclear energy, e.g. with regard to the unresolved problem of final storage.

Danger of a new totalitarian society

► In his latest book "Nexus", Yuval Harari writes:

"If a 21st century totalitarian network succeeds in conquering the world, it may not be controlled by a human dictator, but by a non-human intelligence. Anyone who sees China, Russia or the post-democratic United States as the main threat to a totalitarian nightmare has not understood the danger. The real threat to the Chinese, Russians, Americans and all of humanity is the totalitarian potential of non-human intelligence."

Would you agree with his warning?

Karl Hans Bläsius: Yes, I agree with this warning. It is to be expected that new improved systems of generative AI will be published in the next few years. In particular, successful links between symbolic AI and neural networks could lead to leaps in quality, on the basis of which logical conclusions and the solving of difficult problems could be possible.

On this basis, knowledge structures could be built up, inconsistencies and falsehoods identified and eliminated and logical reasoning improved.

The result could be that these systems are far superior to us in terms of knowledge, judgment and problem-solving skills. These systems could be asked to solve any problem and they would deliver better results than humans would be able to. Human advice would become superfluous, as people would be more likely to trust the results of AI systems than human experts.

This will then also apply to politicians who are also, or possibly even predominantly, guided by the results of these systems, regardless of which party they belong to. Different interests could be represented in different parties, but the AI systems will also be able to justify which of these interests are more important and should therefore be pursued.

The question then arises as to why political parties are still needed and what power politicians have at all if the population can find convincing answers to all problems from AI systems. If differences of opinion between different parties disappear or become irrelevant, different parties will no longer be needed. There would no longer be a democracy; instead, machines would decide and govern.

► What question is not asked or is asked too rarely that you find important with regard to AI and superintelligence?

Karl Hans Bläsius: After the warnings about the risks posed by AI, there was also speculation as to whether an AI system could "press the button", i.e. trigger a nuclear war. Theoretically, this would be possible through actions in cyberspace, for example, but in my opinion this is rather unlikely. However, a comparable risk can arise in other ways.

The military is currently placing great emphasis on the further development of cognitive warfare techniques. One of the aims of cognitive warfare is to influence the enemy and its population, but also its own population, in order to gain support for military plans.

The possibilities for forming opinions and manipulating people have increased significantly thanks to digital media, combining findings from various fields such as psychology, neuroscience and computer science. Cyberspace and AI are playing an increasingly important role here.

Such skills could also be acquired by systems such as ChatGPT. Generative AI systems could thus acquire enormous abilities to manipulate people, including politicians and the military. People could be put under massive pressure to take actions that they do not actually want to take.

Such manipulations could even involve the use of nuclear weapons. So the question is not so much whether an AI will press the button at some point and trigger a nuclear war, but rather whether a human can be manipulated by an AI system in such a way that this human does it.

For all the risks described here, it is not necessary for the AI systems to develop something like consciousness, feelings or free will. The risks exist independently of such considerations.