

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/379821858>

# Artificial Intelligence and a Weapon of Mass Destruction

Article in American Journal of Chemical and Biochemical Engineering · February 2024

DOI: 10.11648/ajcbe.20240801.11

CITATION

1

READS

446

1 author:



**Mamuka Matsaberidze**

Georgian Technical University

37 PUBLICATIONS 112 CITATIONS

SEE PROFILE



## Review Article

# Artificial Intelligence and a Weapon of Mass Destruction

Mamuka Matsaberidze

Department of Chemical and Biological Technologies, Faculty of Chemical Technology and Metallurgy, Georgian Technical University, Tbilisi, Georgia

### Email address:

m.matsaberidze@gtu.ge

### To cite this article:

Mamuka Matsaberidze. (2024). Artificial Intelligence and a Weapon of Mass Destruction. *American Journal of Chemical and Biochemical Engineering*, 8(1), 1-14. <https://doi.org/10.11648/ajcbe.20240801.11>

**Received:** December 8, 2023; **Accepted:** December 23, 2023; **Published:** February 5, 2024

**Abstract:** For global reality, the boundaries between artificial intelligence and software are not clearly defined. Experts in the field of AI are trying not only to understand the nature of intelligence, but also to create intelligent entities, which is why big problems are created within the framework of an unambiguous understanding of the technical regulations of artificial intelligence. Nick Bostrom (*Professor Faculty of Philosophy at Oxford University. He is the founding Director of the Future of Humanity Institute - FHI, a multidisciplinary research center which enables a few exceptional mathematicians, philosophers, and scientists to think carefully about global priorities and big questions for humanity. The FHI is platform to teams working on AI safety, biosecurity, macro strategy, AI policy, the ethics of digital minds, and various other technology or foundational questions*) talks about the need to pay very close attention serious attention to the risk that the actions of an explosively developing artificial superintelligence can lead to the arbitrariness of machines and, as a result, causing irreversible damage to human civilization or even to an existential catastrophe - the death of humanity. If humanity wants to survive the advent of super intelligent machines in its world, then the risks from such realities must be taken into account.

**Keywords:** Artificial Intelligence (AI), Chat GPT, Open AI, Robot Chemist, EU Position on Artificial Intelligence Risks, "Potentially Catastrophic Damage", Henry A. Kissinger

## 1. The Essence of the Problem with AI

Today's world is unimaginable without artificial intelligence. Internet search engines, social networks, "smart" devices, traffic lights, and even public transportation work with the help of artificial intelligence. There are already cars with an autopilot function, unmanned aerial vehicles, and marine vehicles.

It is noteworthy that Ada Lovelace<sup>1</sup>, daughter of George Byron, is called the first programmer. She is the author of the algorithms of the first generation of computing machines [1].

Such counting machines were designed by the British mathematician and engineer - Charles Babbage. It should be

noted here that the first computer appeared during the Second World War and was used in the war to read encrypted messages.

Alan Turing<sup>2</sup> proved that any computation can be done digitally. By conceptualizing such ideas, scientists have assumed that by using technology, it is possible to artificially produce intelligence [2].

To date, there are already so-called a chatbot is a computer programs that can communicate with a human via text. When talking to a sophisticated chatbot, you can talk to the computer program for a few minutes and think it's human. The first such chatbot was created in 1966. Her name was Eliza. Eliza answered questions humanly. People talking to her had a hard time distinguishing whether Eliza was a person or a program.

<sup>1</sup> Ada Lovelace, in full Ada King, countess of Lovelace, original name Augusta Ada Byron, Lady Byron, (born December 10, 1815, Piccadilly Terrace, Middlesex [now in London], England - died November 27, 1852, Marylebone, London), English mathematician, an associate of Charles Babbage, for whose prototype of a digital computer she created a program. She has been called the first computer programmer.

<sup>2</sup> Alan Mathison Turing, (born June 23, 1912, London, England - died June 7, 1954, Wilms low, Cheshire), British mathematician and logician who made major contributions to mathematics, cryptanalysis, logic, philosophy, and mathematical biology and also to the new areas later named computer science, cognitive science, artificial intelligence, and artificial life.

However, Eliza herself had no idea what she was saying, her conversational language was programmed in advance. Eliza was such an important innovation that even a term named after her was established - the Eliza effect<sup>3</sup> [3].

Computers compete with humans in solving various intellectual tasks. In 1997, the computer "Deep Blue" defeated<sup>4</sup> the current world chess champion - Garry Kasparov [4].

Mechanized "creatures" have become a reality today. They perform the most complex calculations, play intellectual games, create technologies, and most importantly, learn.

Artificial intelligence does not have the form that it is often presented to us in science fiction. However, today's world is unimaginable without artificial intelligence. It is used by almost all leading companies, be it Google or Apple. It is artificial intelligence that helps us when searching for information on the Internet, or even when taking pictures with a smartphone. Artificial intelligence is actively used in modern industry. Also, in transport, health, finance, communication, and many other directions, today the new generation of cars already uses a clear form of artificial intelligence, whether it is a voice assistant or a road analysis function. Today's patient sees a robot surgeon and a virtual assistant caring for the patient. Also, an automatic scanner of the tomographic image has already been created, which can make a diagnosis accurately. Modern industry is most dependent on artificial intelligence and automated services. High-quality robots can work faster than humans and cost-effectively over long periods of time.

Artificial intelligence<sup>5</sup> has gone through an impressive path of development and reached unprecedented heights [5]. That is why there is a legitimate question about how reliable artificial intelligence is, how much risk there is to let us say that the personal information on our mobile phones is accessible to a third party?

The International Standardization Organization is also working on these and other challenges of artificial intelligence, which has developed many international standards in the field of personal information protection in the history of its existence, and artificial intelligence is no exception.

That is why, in 2020, the mentioned organization, in cooperation with the International Electronic Commission<sup>6</sup> (IEC) [6], developed a new ISO/IEC TR 24028:2020 standard<sup>7</sup> [7], with the help of which, all "smart" devices that work with the help of artificial intelligence will become much safer and more secure, thus significantly increasing its

reliability. With the help of such a standard, it is possible to analyze all the factors that can damage the reliability of artificial intelligence, and at the same time, the standard can be used by all types of organizations, regardless of industry or size.

Chat GPT [8], a publicly available artificial intelligence program, became popular immediately after its release and became the fastest-growing consumer application in history. It already had more than a million users within the first week of its release on the market. It helps users with various tasks, starting with writing e-mails and literary essays, creating musical or pictorial works, and ending with code generation in various programming languages. Many specialists recognize that Chat GPT<sup>8</sup> is a revolutionary technology, and some realize that it will soon be as well-known as "Google".

Chat GPT's impact is more far-reaching than it might seem at first glance. According to the report of the global investment giant, "Goldman Sachs", which was published [9]. Artificial intelligence already more or less touches about 2/3 of the jobs in the United States. This means the so-called "White-collar" workplaces<sup>9</sup> [10] that will be replaced by artificial intelligence. These can be specialties in the field of jurisprudence or positions in the direction of administrative or human resources management. According to the report, the changes will not affect the service sector and such workplaces where physical labor is required.

The first version of Chat GPT released in late November 2022 was free. Currently, it has more than 100 million monthly active users. In February 2023, its paid counterpart Chat GPT Plus was released.

Both programs were developed in the American research laboratory Open AI ("Open AI") - it translates as "open artificial intelligence". The colossal popularity of Chat GPT prompted the technology giants to quickly start work on creating similar programs.

Among them is the company "Microsoft", which launched "Bing-Chat", according to reports, a similar project is being developed at "Facebook". The implementation of artificial intelligence systems in the public sector of many countries is at the initial stage of development, however, in the private sector there are many successful examples of the use of this technology, remote verification systems, automatic document identification systems, communication automation programs, and many other tools. However, the risks [11] that may arise as a result of system abuse, technical errors, and incorrect management of personal data are also to be considered.

Artificial intelligence is not just another electronic assistive system. It essentially increases the governmental capacity of the state, therefore, the temptation to use it irregularly increases. This risk is particularly high in developing

3 ELIZA, the first chatbot program, developed in the 1960s at the Massachusetts Institute of Technology by Joseph Weizenbaum to engage in written conversations with users of the MAC time-sharing system.

4 Garry Kasparov lost two games to one, with three draws.

5 The AI Index Report tracks, collates, distills, and visualizes data related to artificial intelligence. Our mission is to provide unbiased, rigorously vetted, broadly sourced data in order for policymakers, researchers, executives, journalists, and the general public to develop a more thorough and nuanced understanding of the complex field of AI.

6 The IEC is a global, not-for-profit membership organization that brings together more than 170 countries and coordinates the work of 20 000 experts globally.

7 ISO/IEC TR 24028:2020 standard: Information technology - Artificial intelligence - Overview of trustworthiness in artificial intelligence.

8 Chat Generative Pre-Trained Transformer, is a large language model-based chatbot developed by Open AI and launched on November 30, 2022, notable for enabling users to refine and steer a conversation towards a desired length, format, style, level of detail, and language used.

9 A "white-collar" worker belongs to a class of employees known for earning higher average salaries doing highly skilled work, but not by performing manual labor at their jobs.

democracies. Research has established that law enforcement agencies [12] are the only area of the public sector where the process of introducing artificial intelligence is going on steadily, which indirectly indicates<sup>10</sup> the reality of these risks<sup>11</sup>. The AI Index Report tracks (Artificial Intelligence Index Report 2023 of Stanford University), collates, distills, and visualizes data related to artificial intelligence. Their mission is to provide unbiased, rigorously vetted, broadly sourced data in order for policymakers, researchers, executives, journalists, and the general public to develop a more thorough and nuanced understanding of the complex field of AI. The report aims to be the world's most credible and authoritative source for data and insights about AI [13].

The study, published in November 2023 in the journal *Nature Synthesis*, details the steps taken by the robot chemist [14] in creating a way to extract oxygen from materials available on Mars. Using meteorites believed to have originated from the Red Planet, a robot chemist was able to create the primary element needed to support human life. Oxygen supply is becoming a key priority for future human activities on Mars, as rocket fuel and life support systems require large amounts of oxygen, which cannot be obtained from the Martian atmosphere, according to the study. Recent discoveries of signs of aquatic activity open the prospect of large-scale oxygen production on the planet using electrochemical processes of water oxidation under the influence of solar energy using an oxygen evolution reaction (OER) catalyst. The use of extraterrestrial catalysts based on local materials for the production of oxygen provides the prospect of producing fuel and oxygen on Mars, which is one of the promising areas of planetary research. The possibility of reaching Mars remains one of the main goals of those who strive for further exploration of the solar system. In the summer of 2023, the National Aeronautics and Space Administration (NASA) launched a test simulating life in a bio dome.

## 2. About the Open Letter on the Risks of Uncontrolled Development of Artificial Intelligence

Some specialists in the field think that the uncontrolled development of artificial intelligence contains serious risks. "Artificial intelligence is more dangerous than the mishandling of an airplane and its maintenance, or a poorly manufactured car". Billionaire businessman Elon Musk, in a recent interview, claimed that artificial intelligence (AI) has the potential to cause "civilization destruction"<sup>12</sup>. His remarks

come even as he is deeply invested in AI technology and is all set to come out with an AI platform, Truth GPT, "a maximum truth-seeking AI that tries to understand the nature of the universe" [15]. Elon Musk is one of the signatories of the high-profile open letter, which calls for the suspension of the development of artificial intelligence for at least 6 months.

The pause should be public, transparent, and controlled, supported by an international intergovernmental memorandum [16] - say the initiators of the open appeal from the "Future of Life Institute".

According to their argument, the suspension is necessary so that work on regulating artificial intelligence can begin as quickly as possible.

The purpose of this should be to minimize risks and ensure the safe operation of programs. "As intensive research has shown and recognized by leading AI laboratories, artificial intelligence systems that compete with the human mind can pose profound risks to society and humanity..."

The development of strong artificial intelligence<sup>13</sup> systems should only happen when we are convinced that their results will be positive, and their risks are manageable" - it is written in the appeal [17]. In addition to Elon Musk, thousands of people have signed the public appeal. Among them is such a well-known face in the field of technology as Apple founder Steve Wozniak.

The group appeal is also signed by professors from the Massachusetts Institute of Technology, Harvard, Cambridge, Yale, Carnegie Mellon, New York, Oxford, and many other highly ranked universities.

EU position on artificial intelligence risks - the European Union and Canada have already started looking into how Open AI collects private information and then uses it in the hugely popular Chat GPT.

Italy became the first country in the West where it was temporarily banned in early April 2023. Ireland has given Chat GPT until April 30, 2023, to comply with current data protection regulations, which industry experts say is a virtually impossible task.

Europol, the European Union's police service, has issued a warning about the AI software, saying its use could make it easier for would-be cybercriminals and those seeking to spread disinformation.

"The possible use of this type of artificial intelligence systems by criminals is a dark prospect," - states Europol in a special report [18].

The situation in the USA - It is clear that the United States government is also considering the issue of developing regulations. Both the executive and legislative branches of government are involved in the discussions: the White House, the Federal Trade Commission, and Congress.

The U.S. Department of Commerce's Telecommunications and Information Administration agency also issued a

10 The INTERPOL Face Recognition System (IFRS) contains facial images received from more than 179 countries which makes it a unique global criminal database.

11 AI could cause nuclear-level disaster, third of experts tell poll Stanford University report says "incidents and controversies" associated with AI have increased 26 times in a decade //

<https://www.aljazeera.com/economy/2023/4/14/ai-could-cause-nuclear-level-catastrophe-third-of-experts-say>

12 AI more dangerous than mismanaged aircraft design; can destroy civilization //

[https://www.theweek.in/news/biz-tech/2023/04/18/ai-more-dangerous-mismanage-d-aircraft-design-can-destroy-civilization-elon-musk-views-artificial-intelligence-](https://www.theweek.in/news/biz-tech/2023/04/18/ai-more-dangerous-mismanage-d-aircraft-design-can-destroy-civilization-elon-musk-views-artificial-intelligence-troubles.html)

[troubles.html](https://www.theweek.in/news/biz-tech/2023/04/18/ai-more-dangerous-mismanage-d-aircraft-design-can-destroy-civilization-elon-musk-views-artificial-intelligence-troubles.html)

13 Strong artificial intelligence (strong AI) is an artificial intelligence construct that has mental capabilities and functions that mimic the human brain //

<https://www.techopedia.com/definition/31622/strong-artificial-intelligence-strong-ai>

statement saying there is "growing regulatory interest" in creating an "accountability mechanism" for AI developments.

"Time is running out to get ahead of this powerful new technology and prevent potentially large-scale harm to society and national security." And instead, let's use it on the positive side by advancing strong, bipartisan legislation," Senate Majority Leader Chuck Schumer said as he introduced the Chat GPT regulations working framework document on the Senate floor.

According to Senator Schumer<sup>14</sup>, the framework document [19] is aimed at creating new legal regulations that will prevent "potentially catastrophic damage" in the United States, but at the same time, it will become a prerequisite for the development and promotion of such revolutionary technology in America.

Chat GPT's representatives regarding the initiative to take a 6-month hiatus say that some of the ideas put forward in the collective appeal are perfectly acceptable and they have been working hard to ensure the security of the systems for some time. As for government regulation, "Yes. These systems need to be regulated," says Mira Murat<sup>15</sup>, Chief Technology Officer of Chat GPT, whose name is associated with Chat GPT's transformation and success in recent years?

M. Murat believes that "at some level" standards are needed, and they are systematically consulting on the issue with governments, regulators, and other organizations working to develop similar systems. However, in his opinion, "much more should be done. Government regulators should be widely involved," she said in an interview with The Associated Press [20].

Artificial intelligence is a "more urgent" threat to humanity than climate change -

Geoffrey Hinton [21], who has been called the "father of artificial intelligence", has said that the rise of artificial intelligence (AI) could pose a "more urgent" threat to humanity than climate change. "I don't want to devalue climate change," Geoffrey Hinton, 75, told Reuters, adding: "I don't want to say don't worry about climate change, which is too big a risk, but I think it might become more relevant." According to him, the decision was made so that the artificial intelligence, without the supervision of his former employer could create a platform where he would be able to issue warnings on the issue of risks.

Google has launched an AI chatbot 'Bard' to compete with Open AI's Chat GPT software.

Hinton's warning comes after several tech industry leaders, including SpaceX, Tesla, Elon Musk signed an open letter warning about the potential consequences of an artificial intelligence arms race.

### 3. Henry Kissinger's Article "The End of the Age of Education"

In the American magazine The Atlantic, in 2018, Henry

Kissinger's [22] article "The End of the Education Era" [23] was published, which caused a global resonance. H. Kissinger<sup>16</sup> is interested in such advances in artificial intelligence as the creation of the Alfa Go<sup>17</sup> game-winning algorithm [24], which is not the result of programming.

In this case, the artificial intelligence replayed the games with itself many times to learn from its own mistakes and get the perfect winning algorithm. This resulted in the creation of Alfa Go, an artificial intelligence program equipped with the ability to beat world leaders in the Chinese board strategy game Go, which is more difficult than chess.

The AlphaGo project was launched in 2014 as a testbed to see how well Google's DeepMind neural network algorithm [25], which uses deep learning, can compete with Go. The AlphaGo algorithm [26] is a combination of tree search and machine learning techniques, backed up by extensive training from both humans and other computer players. It uses a Monte Carlo tree search and is guided by a policy and value network implemented using deep neural network technologies.

The policy network is trained to help the AI predict the next move most likely to win, while the value network is trained to narrow down the search tree and determine the value of those positions by ranking the winners in each position rather than searching everything down. until the end of the game.

AlphaGo first loaded historical match moves from human players using a database of around 30 million moves, allowing it to simulate human play. Once the AI reached a certain degree of mastery, it was taught further by having it play against instances of itself, using reinforcement learning to improve and learn more.

In October 2015, the distributed computing version of AlphaGo played and defeated Fang Hui, a 2-Dan European Go champion, marking the first time a computer program had beaten a professional Go player. A few months after his defeat, Fang Hui helped the DeepMind team as a consultant. In March 2016, AlphaGo met Lee Sedol, one of the top 9-dan players in the world.

Winning four games against Li alone, was a breakthrough in AI research as it meant that the deep learning and neural network algorithm used by DeepMind could be used for any other purpose since it had not been programmed to play Go. but rather was trained. how to play go. This opens up a whole new world for AI research.

Henry Kissinger<sup>18</sup>, as a historian and public figure, asked

16 Henry A. Kissinger, the scholar-turned-diplomat who engineered the United States' opening to China, negotiated its exit from Vietnam, and used cunning, ambition and intellect to remake American power relationships with the Soviet Union at the height of the Cold War, sometimes trampling on democratic values to do so, died on Wednesday at his home in Kent, Conn. He was 100. <https://www.nytimes.com/2023/11/29/us/henry-kissinger-dead.html>

17 a computer program developed by Google DeepMind to play Go, a Chinese strategy board game for two players similar to chess. AlphaGo is the very first AI program that was able to beat a professional human player, 2-dan player Fan Hui in October 2015, on a full-sized board with no handicaps. It then beat one of the highest ranked human players in the world, 9-dan Lee Sedol, in March 2016, winning four games out of five.

18 (born Heinz Alfred Kissinger) is a German-born American bureaucrat, diplomat, and 1973 Nobel Peace Prize laureate.

14 US Senate leader Schumer.

15 Chief technology officer, leads Open AI's research, product and safety teams.

questions about how artificial intelligence will affect human history. What will the self-educated machines that have acquired knowledge through specific internal processes and can use these achievements for various purposes, which can no longer be controlled by humans, this was possible in the example of the AlphaGo game. In connection with this, the following questions are relevant:

- 1) Will machines learn to communicate with each other?
- 2) Is humanity facing a new historical stage?

To this day, the "digital" breakthrough that created modern history is considered to be writing, and its technical implementation is the typewriter. Knowledge as a result became available to the masses. Human scientific knowledge appeared as one of the main criteria of consciousness.

Information was stored and systematized in man-made and intensively growing libraries. The realization of scientific knowledge shaped the modern world order. But today, amid the rapid digital revolution, this order is transforming, the consequences of which are currently unpredictable. As a result, a different world can be created, establishing a completely different axiological system. It is noteworthy that such a prominent scientist and statesman as Henry Kissinger appealed to scientists working in different fields to unite their efforts to understand what results the development of artificial intelligence will bring in the future and what impact it will have on human evolution.

First of all, it should be noted that the game is created by the given example of Alfa Go to a specific question regarding the determination of the level of artificial intelligence. As we have already mentioned, human intelligence is a very complex system and its main function is to perceive the world with gestalt [27].

The following question can be asked: do the results obtained in the Alfa Go game experiment indicate the ability of artificial intelligence to create a new whole, such as Gestalt, in the process of self-learning? Gestalt describes a special type of systemic integrity, in which the main characteristic feature is that the mentioned properties of the whole exceed and differ from the properties of the parts that make up this whole. So, the union produces a qualitatively different phenomenon.

Will artificial intelligence, as a mathematical algorithm, be able to create a new systemic entity, the properties of which were not originally established? There is one of the main paradigms in the program to evaluate the artificial perspectives on intelligence this issue is important both theoretically and theoretically from a practical point of view [28].

The knowledge we have about human evolution does not allow us to determine exactly under what conditions human intelligence was formed during the evolution process. Therefore, when we create artificial intelligence, we cannot in principle create something that corresponds to human intelligence. It is a leap into the unknown that requires great caution. From this point of view, we believe that a deeper understanding of the phenomenon of ambivalence as one of the characteristics of the human psyche is necessary.

The main danger is that machines equipped with artificial

intelligence are created, which have the ability to learn independently, to solve tasks without human intervention, but you cannot teach any machine to express emotion, to be guided by morality, to define exceptions, to absorb religious and racial doctrines.

Artificial intelligence is not subject to human feelings [29], and therein lies the whole problem<sup>19</sup>.

No one knows what conclusions a robot equipped with artificial intelligence will reach in the end.

H. A. Kissinger is a specialist in nuclear weapons issues. In 1957, he published the book "Nuclear Weapons and Foreign Policy", thanks to which he became known throughout the United States. In the pages of this book, he, at that time a professor at Harvard, explained how, with the help of local nuclear strikes, victory in the war could be achieved.

Since then, the views of the scientist have changed and H. A. Kissinger is one of the founders of the Global Zero civil organization, whose goal is the destruction of all nuclear weapons by 2030. But despite this, H. A. Kissinger insists that the United States is obliged to lead in the field of nuclear technologies, regardless of the negotiation on the impossibility of using this type of weapon.

Scientist considers cybersecurity to be another scientific and technical challenge. H. Kissinger also reminded me about computer viruses. The "Stuxnet" virus, an example of an Israeli government-backed cyberattack, brought Iran's nuclear program to a halt, and the damage was proportional to a limited nuclear strike. And another virus, known as "Russian Bot", left Estonia without power for several days.

The scientist emphasized that today a person with a smartphone has information and analytical power that were not available to most intelligence agencies just a generation ago, and this power is growing at an unimaginable rate. The absence of a strategy of behavior and an idea of what it should be can lead to world chaos and not order.

H. Kissinger also returns to the issue of artificial intelligence in an article devoted to the victory of the AlphaGo algorithm over world go champion Lee Sedol [Kissinger, H. A. How the Enlightenment Ends / H. A. Kissinger // Journal the Atlantic. 06/01/2018. p. 12-16].

The scientist drew attention to the fact that go is more difficult for artificial intelligence than classical chess, because, in addition to intuition, it requires the use of creative and strategic thinking. Therefore, H. Kissinger reflects on who can use such power in the future, how this will affect the global balance of power, and also how to establish a balance between nation-states if the latest achievements are used by non-state organizations that do not seek to protect a certain territory and reject established principles of legality.

The leader of geopolitics - Henry Kissinger, in a brilliant article published in 2018 (How the Enlightenment ends philosophically, and intellectually - in every way - human society is unprepared for the rise of artificial intelligence.

19 In Moscow the robot did not like the child's haste; he grabbed the boy's index finger and squeezed it hard. The Chess robot broke the finger of 7-year-old Christopher, who is in the top 30 strongest chess players in Moscow under 9 years old.

Henry Kissinger warns modern society about the risks [30] of unpredictability in the field of artificial intelligence: “Three years ago, at a conference on transatlantic issues, the topic of artificial intelligence (AI) appeared on the agenda. I was about to leave the meeting as this subject was not my area of interest, but the start of the presentation forced me to stay. The speaker talked about a computer program that could soon challenge world champions in the game of “Go”<sup>20</sup> [31]. I was surprised: can a computer handle “Go” – a more complex game than chess? Each player has 180 or 181 stones at his disposal (depending on the chosen color) and places them in turn on an empty board. The winner is the one who, by making effective strategic decisions, paralyzes the actions of the enemy and will control a large territory.

The speaker argued that this ability could not be programmed. His machine mastered the game through practice. Knowing the rules, the computer played countless games against itself and learned from its own mistakes, improving the algorithms. In the process of training, he managed to surpass his human mentors. A few months later, the AlphaGo AI program appeared, capable of beating the strongest “Go” masters.

As I listened to claims of the victory of technological progress, my experience as a historian and occasional statesman made me think. How will history be affected by self-learning machines—those that acquire knowledge through specific internal processes and can apply it to purposes beyond human understanding? Will machines learn to communicate with each other? How will you select from the available options? Can the fate of the Incas, who faced the Spanish culture, which was inaccessible to their understanding and caused horror, be repeated in the history of mankind?

Are we facing a new phase of history? Realizing my incompetence in this area, I organized a series of informal meetings with specialists in technology and the humanities. The discussions that took place disturbed me even more.

Until now, the technological breakthrough that changed the course of modern history could be called the invention of the printing press in the 15th century. As a result, access to empirical knowledge appeared, which supplanted religious doctrines; the era of religion was gradually replaced by the era of reason.

Human insight and scientific knowledge have replaced faith as the main criterion of consciousness. Information was stored and systematized in growing libraries. The Age of Reason stimulated the thoughts and actions that shaped the modern world order. But today this order is undergoing a transformation against the backdrop of a new, even faster technological revolution, the consequences of which we do not yet understand. The result could be a world based on machines driven by data and algorithms rather than ethical and philosophical norms. We live in the age of the Internet and are already facing some issues that will become even more acute

with the spread of artificial intelligence. The Enlightenment exposed traditional truths to the free, analytical human mind.

The purpose of the internet is to validate knowledge through the accumulation and manipulation of ever-increasing amounts of data. The cognitive abilities of a person lose their character. The individual becomes a set of data, and the data begins to dominate.

In the digital world, speed is a priority, which hinders reflection, radicals get an advantage over thoughtful people, and values are formed by subgroup consensus, and not in the process of reflection. Despite all the achievements, the digital world is in danger of destroying itself, because there are more disadvantages than amenities.

The Internet and the increasing power of computers have facilitated the accumulation and analysis of data, and unprecedented resources have opened up that are difficult for a person to comprehend. One of the most significant is artificial intelligence, i.e. a technology that allows us to solve complex, seemingly abstract problems using processes that resemble the work of the human brain. It's not just about automation, as we now understand it. Automation refers to the means to achieve goals by rationalizing or mechanizing tools. Artificial intelligence, on the contrary, is related precisely to goals, it sets its tasks. Since its achievements are partly determined by itself, artificial intelligence is inherently unstable. AI systems in the process of functioning are constantly changing: they receive and instantly analyze new data, and then improve themselves based on this analysis.

As a result, artificial intelligence acquires an ability that was previously considered the prerogative of a person. He makes strategic judgments about the future based on set data (such as the rules of the game) or data he has obtained (such as playing 1 million games).

The self-driving car demonstrates the differences between traditional human-controlled and software-defined computers and a universe where artificial intelligence will reign. When driving a car, it is necessary to evaluate many situations that cannot be foreseen and therefore cannot be programmed.

Take a well-known hypothetical example: what happens if a self-driving car has to choose between the death of a grandmother and the death of a child? Who will he choose? Why? What factors will try to optimize? Will he be able to explain his actions? If he could communicate, he would most likely respond like this: “I don’t know (because I follow mathematical, not human principles)” or “You will not understand (because I am trained to act in a certain way, and not explain my actions)”. Meanwhile, in 10 years, unmanned vehicles will dominate the roads.

Until now, AI research has been limited to certain areas of activity, but now it is about developing “widely developed intelligence” capable of performing tasks in various areas.

Shortly, AI algorithms will work in many areas of human activity. However, they are a mathematical interpretation of the observed data and do not explain the realities that caused them.

Paradoxically: The world becomes more transparent and at the same time more mysterious. How will the new world be

20 The game was invented in China more than 2,500 years ago and is believed to be the oldest board game continuously played to the present day.

different from what we are used to today? How will we live in it? How will we manage artificial intelligence, improve it, and in extreme cases, prevent it from harming us? The main problem is that artificial intelligence learns skills faster and more confidently than a person, so over time it can reduce human activity and turn people into data.

Artificial intelligence can bring incredible benefits in medicine, clean energy, ecology, and many other areas. But precisely because AI makes decisions about an as yet uncertain future, the results of its activities are inherently uncertain and ambiguous".

Henry Kissinger believes that three aspects are of particular concern:

First, artificial intelligence can achieve unintended results. Science fiction describes how artificial intelligence attacks its creators. It is more likely that the AI misinterprets human instructions due to a lack of understanding of the context. As an example, we can recall the Tay [32] AI virtual interlocutor, which is designed to conduct friendly conversations in the language of a 19-year-old girl. But the machine could not recognize the imperatives "friendly" and "reasonable" set by the developers, so the virtual interlocutor gave racist, sexist and simply offensive remarks. Some in the tech world argue that the experiment was ill-conceived and poorly implemented, but it nevertheless demonstrated basic ambiguity. To what extent is the AI able to understand the context that defines the instructions it receives? Who would help Tay understand the meaning of the word "aggressive" if the people themselves cannot agree? Can we spot bugs early on and correct an AI program that doesn't work the way we expected? Or will artificial intelligence inevitably have small deviations, which over time can lead to disastrous consequences?

Secondly, in the process of achieving the set goals, AI can change human thinking and human values. AlphaGo managed to beat the strongest Go masters by taking unprecedented strategic steps that man could not even think of and which he has not yet learned to counteract. Are these moves beyond the capabilities of the human brain? Or will people be able to master them after being shown by a new go master?

Before artificial intelligence started playing Go, the game had various multi-level tasks. The player sought not only to win, but also to master new strategies applicable in other life situations. Artificial intelligence has one goal - to win. He "learns" not conceptually, but mathematically, improving his algorithms. The AI has learned to win by playing differently than a human, and has succeeded in changing the nature of the game and its impact. Can we say that such a one-sided approach is characteristic of artificial intelligence in general?

Other AI projects are engaged in the modification of human thinking by developing devices that can generate responses to human requests. In addition to questions about facts (what is the temperature outside?), questions about the nature of reality and the meaning of life raise deeper issues. Do we want children to learn about values through dispassionate algorithms? Should we protect privacy by restricting AI information about those who ask questions? If yes, how are we going to do it? If artificial intelligence learns many times

faster than a person, we can expect that the process of trial and error (as people usually make decisions) will also go faster for him. Only AI mistakes will be made faster and with more serious consequences. It is unlikely that these errors can be dealt with by adding an "ethical" or "rational" calculation of the situation to the program, as suggested by AI developers. All scientific disciplines have arisen because of the inability of mankind to agree on the definition of these terms. So, can artificial intelligence be the arbiter?

Third, the AI may achieve its intended goals but fail to explain its findings. In some areas – pattern recognition, big data analysis, games - AI is already superior in its abilities to humans. If its computing resources continue to grow as rapidly, AI will soon be able to optimize situations, but not in the way that a human would. But at this stage, can AI explain why its actions are optimal, so that a person can understand? Or are AI decisions defying human language and therefore inaccessible to human understanding?

Throughout human history, civilizations have created ways of explaining the world: in the middle Ages it was religion, in the Enlightenment it was reason, in the 19th century it was history, in the 20th century it was ideology. But the most difficult and important question about the world we are moving towards is:

- 1) What will happen to the human mind if its explanatory abilities will be inferior to artificial intelligence and society will no longer be able to interpret the existing world in terms that make sense to it?
- 2) How to define consciousness in the world of machines that limit human experience to mathematical data interpreted by their own memory?
- 3) Who is responsible for AI actions?
- 4) How will responsibility for his mistakes be established?
- 5) Will a human-made legal system be able to control the activities of an AI that is smarter and potentially more cunning than a human?

Ultimately, the term "artificial intelligence" runs the risk of being misleading. These machines can indeed solve complex, seemingly abstract problems that were previously only amenable to human cognition. But what is their unique feature is that they do not think on the basis of the knowledge and experience that they have previously acquired. Rather, it is about unprecedented memory capabilities and computing resources. Given the natural superiority of AI in these aspects, it will always win games. But for humanity, games are not only a victory, reflections are important for us.

If we treat the mathematical process as a thought process, try to imitate it and accept its results unconditionally, then we are in danger - we can lose the ability that is the basis of human knowledge.

The problems of this evolution are demonstrated by the recently developed program Alpha Zero<sup>21</sup> [33], which plays chess at a level above grandmaster and in a style not seen

---

21 This algorithm uses an approach similar to AlphaGo Zero developed by artificial intelligence research company DeepMind to master the games of chess, shogi, and go.



before in the history of chess. In a few hours of playing with yourself, the program reached a level that took humanity 1,500 years to achieve. Only the basic rules of the game were loaded into the program. Humans and the data they collect were not used in Alpha Zero's self-learning process. If the program was able to achieve this level of proficiency so quickly, where will artificial intelligence be in five years? How will this affect human cognitive abilities? What is the role of ethics in this process, which is essentially a fast-track choice?

Usually such questions are left to technologists and intellectuals from related scientific fields. Philosophers and other humanists who helped formulate the concepts of the world order do not enter into the discussion because they lack knowledge about the mechanisms of AI or its possibilities terrify them.

The scientific world, on the contrary, is ready to explore the technical possibilities of its achievements, while the technological world is busy with large-scale commercial implementation of its ideas. Both worlds seek to push the boundaries of discovery without understanding them. And the authorities are more interested in the use of AI in the field of security and intelligence than in the already begun transformation of human life.

The Enlightenment began with philosophical reflections that were spread through new technology. Our era is moving in the opposite direction. A potentially dominant technology has been developed that needs a guiding philosophy. In many countries, AI has become a national project. The United States has not yet systematically explored the full range of AI capabilities, its impact, and has not begun the process of improvement. This should become a national priority in terms of the relationship between AI and humanist traditions.

AI creators who are as incompetent in politics and philosophy as I (i.e. Henry Kissinger) are in the field of technology should ask the questions I have raised in this article in order to build the answers into their engineering developments. "It is known that the opportunity to benefit society with the help of artificial intelligence technologies in the short-term drive's developers in many areas. Modern developments are used to confirm economic theories, in lawmaking, technical developments in the field of verification, data updating, security and control. So, if a hack or malfunction in a personal computer system can go unnoticed, malfunctions in the work of an AI<sup>22</sup> that controls a car, an airplane, a pacemaker, an electrical grid, or, for example, an automated trading system, can lead to much larger and more tragic consequences. In addition, the need to prevent a destructive arms race in the field of autonomous weapons [34] is becoming increasingly urgent today.

What will happen if AI surpasses the human mind and can more successfully solve all cognitive tasks? In the mid-60s of

the twentieth century, Irving John Good<sup>23</sup> [35] noted that the improvement of the AI system is in itself a cognitive task. Such a system is potentially capable of recursive self-improvement and can thus provoke an "explosion" of intelligence, leaving the human mind far behind.

The invention of revolutionary technologies, and superintelligence is one of them, can help humanity eradicate war, disease and poverty. In itself, the creation of such technology could be the biggest event in the history of mankind.

However, some experts suggest that if a person does not learn to synchronize his goals with the tasks of AI before he turns into a superintelligence, this event may be his last. There are those who doubt the possibility of creating strong AI, and those who believe that the creation of super advanced AI is guaranteed to benefit humanity.

The Future of Life Institute [36] recognizes the possibility of both scenarios and understands that AI can, intentionally or unintentionally, become a source of great misfortune. We believe that structured research today will help prepare for and prevent potential negative scenarios in the future, and will provide an opportunity to avoid the dangers associated with AI, and instead use its advantages.

Most researchers agree that a super AI is unlikely to display human emotions, such as love or hate. Nor should one expect artificial intelligence to become deliberately benevolent or, conversely, embittered. According to experts, the most dangerous AI systems can be the result of two scenarios:

I. Artificial intelligence was originally developed with a destructive purpose: for example, as a control system for autonomous weapons. In the hands of the wrong person, such a weapon would easily cause many casualties. What's more, an arms race using AI technology could inadvertently lead to AI wars that will undoubtedly claim even more lives. Most likely, such a weapon will not be so easy to take and "turn off" - systems will be developed to protect against enemy hacking of systems. And as a result, humanity may lose control over such weapons. There is a risk even when using weak AI, but with an increase in the degree of its intelligence and autonomy, the threat increases many times over.

II. Artificial intelligence was originally designed for a positive task, but in the process of its execution it chooses a destructive method of achieving the goal: this can happen as soon as we make a mistake in synchronizing the goals of artificial intelligence and humanity. Synchronization is incredibly complex. So, if you ask a self-driving car to take you to the airport as soon as possible, you can arrive at your destination with a motorcade of police chasing you - the trip will be completely uncomfortable. The machine will literally do what you asked, but not what you really wanted. If the superintelligence is entrusted with an ambitious geoengineering project, the side effect of its activities may be

22 AI has reached a point where the deployment of such systems is - practically if not legally - feasible within years, not decades, and the stakes are high: autonomous weapons have been described as the third revolution in warfare, after gunpowder and nuclear arms.

23 Irving John Good (9 December 1916 – 5 April 2009) was a British mathematician who worked as a cryptologist at Bletchley Park with Alan Turing. After the Second World War, Good continued to work with Turing on the design of computers and Bayesian statistics at the University of Manchester. Good moved to the United States where he was professor at Virginia Tech.

the destruction of the ecosystem, and human attempts to stop the work of AI will be perceived as a threat that must be repelled.

All these examples show that the main difficulty in applying strong AI systems is not related to evil superintelligence, but rather to the competence of the developers. Super AI will be incredibly effective in solving the tasks, but if they disperse from the real goal of a person, problems will arise. A mentally healthy person does not consider himself an ant exterminator who, out of pleasure of a sick psyche and passion, tramples insects, but if people are faced with the task of building a hydroelectric power station, and an anthill accidentally appears in the flood zone, so much the worse for ants. A key goal of AI security research is to prevent a situation in which humanity will be in the place of ants.

The idea of creating strong AI has always been considered science fiction, and it seemed to take more than one century to implement it. Despite this, thanks to recent breakthroughs, many key milestones, the time to which scientists until recently estimated in decades, have already been passed. Now many experts are seriously thinking about the possibility of creating a superintelligence during the life of the modern generation. While some believe that there is still a century before the creation of human equal AI, at a conference<sup>24</sup> in Puerto Rico [37] in 2015, most researchers agreed that superintelligence will appear before 2060. Since the creation of security protocols can take decades, it is wise to start acting now.

AI has the potential to surpass the cognitive abilities of any human, so there is no sure way to predict how the system will behave. We cannot use the developments of the past as a basis for predictions, since humanity has never before created mechanisms that can accidentally or deliberately become smarter than us. The best example of what we can face is the evolution of man himself. Humans dominate this planet not because they are stronger, larger or faster than other species, but because they are smarter. And if we become not the smartest, how can we maintain dominance? The Future of Life Institute believes that civilization can flourish as long as we don't allow the growing power of technology to take precedence over its prudent use. As far as AI technologies are concerned, we think the key is not to hinder their development, but to focus on their wise use through research into AI security.

According to leading experts in the field, research on artificial intelligence (AI) from the very beginning included the study of a wide range of problems and directions, over the past decades, special attention has been paid to the problems associated with the construction of intelligent agents - systems that perceive the environment and act in it. In this context, "intelligence" is associated with the statistical and economic

concepts of rationality, in other words, with the ability to make the right decisions, plan or draw conclusions.

The application of probabilistic representations and decision theory, as well as the use of statistical learning methods, has led to significant integration and cross-fertilization of AI, machine learning, statistics, control theory, neurosciences and other fields. The establishment of common theoretical foundations, combined with access to data and computing power, has led to significant advances in solving various complex problems, such as speech recognition, image classification, machine translation, walking simulation, autonomous vehicles and question-answer systems.

As the commercial potential of these and other applications of AI technologies expands beyond purely laboratory research and acquires economic value, a virtuous cycle of technology development is emerging, where even small improvements in functions lead to a noticeable increase in economic value, attracting investors.

There is a widespread belief today that AI research is steadily advancing and that its impact on society is likely to continue to grow. The potential benefits of using AI are enormous, given that all the achievements of modern civilization are the product of human intelligence, and it is difficult to imagine what other heights can be achieved by enhancing the human mind with AI tools, but even the eradication of disease and poverty seems possible.

The potential of AI is enormous, and therefore it is important to understand how to take advantage of it and not fall into a dangerous trap. Progress in the development of AI dictates an urgent need not only to improve its capabilities, but also to maximize the benefits that AI will bring to society.

To address these concerns, the Association for the Advancement of AI (AAAI) Presidential Expert Group on the Long-Term Perspectives of AI was established in 2008 and other AI impact research projects were initiated, signaling a broadening of the range of AI research, until recently focused on technology rather than technology on target. We support extended research to ensure the reliability and usefulness of AI systems with ever-increasing capabilities: our AI systems should do what we want them to do.

The attached paper on research priorities [38] provides a number of examples of research areas focused on maximizing the societal benefit of using AI. Such research is inevitably interdisciplinary in nature, as it affects both AI and society, and includes a range of disciplines from economics, philosophy and law to computer security, formal methods and, of course, various areas of AI itself. Thus, AI experts<sup>25</sup> [39] believe that scientific research into the problems of ensuring the reliability and usefulness of AI systems [40] is important and timely, and there are specific areas<sup>26</sup> of research that can

24 This conference brought together the world's leading AI builders from academia and industry to engage with each other and experts in economics, law and ethics. The goal was to identify promising research directions that can help maximize the future benefits of AI while avoiding pitfalls (see this open letter and this list of research priorities - opportunities and challenges).

25 An Open Letter: There is now a broad consensus that AI research is progressing steadily, and that its impact on society is likely to increase. The potential benefits are huge, since everything that civilization has to offer is a product of human intelligence. Because of the great potential of AI, it is important to research how to reap its benefits while avoiding potential pitfalls.

26 Artificial intelligence (AI) is a rapidly growing field of technology with potentially significant implications for national security. As such, the United States

be pursued today. If we do not waste time on human delusions, we can concentrate on really interesting points of contention, in which even experts disagree with each other and which can become the subject of intercultural and inter-civilizational dialogue.

One of the questionnaires to discuss during the AI Dialogue might look like this:

- 1) How do you see your future?
- 2) Should we develop autonomous weapons?
- 3) How do you feel about workplace automation?
- 4) What profession would you advise today's children to choose?
- 5) Would you rather see new jobs replace the old ones, or would a society that is free from the need to work blissfully spend its free time enjoying the results of the work of machines?
- 6) If we talk about the distant future, do we want to create super intelligent life and spread it in the Universe?
- 7) Will man control machines or will they take control of man?
- 8) Will intelligent machines replace humanity, will they coexist with it, or will they completely assimilate?
- 9) What will people be like in the era of artificial intelligence?
- 10) What would you like to see this era, and how to make the vision a reality in the future?

The American "AI Security Center"<sup>27</sup> [41] released an official statement. It details a number of negative effects of the use of AI, which could eventually lead to the extinction of humanity. "More advanced AI could pose catastrophic and existential risks" [42].

Risks associated with the military application of artificial intelligence technologies; The AI race [43]: Competition can push countries and corporations to accelerate AI development<sup>28</sup> by ceding control to these systems. Conflicts could spiral out of control with autonomous weapons and AI cyber warfare. Corporations will face incentives to automate human labor, which could lead to mass unemployment and dependence on AI systems.

As AI systems proliferate, evolutionary dynamics suggest that they will become more difficult to manage. We recommend safety regulations, international coordination and public oversight of general-purpose AI:

- 1) a sharp increase in people's dependence on advanced systems up to delegating the most common tasks to

them;

- 2) a more intense concentration of wealth and influence in the hands of a limited number of people who can use this to monopolize power in society;
- 3) Emergence of goals and abilities. In other words, models can set new goals and develop skills that were not foreseen during their training;
- 4) AI cheating.
- 5) the statement cites an example where Volkswagen programmed engines to show low emissions only during inspections.
- 6) in the same way, according to experts, artificial intelligence can do;
- 7) the biggest threat is the desire for power.

According to CAIS, the risks of artificial intelligence aiming to become more independent cannot be ruled out: "In inventing more advanced machines, we are playing with fire."

CAIS experts conclude: "Reducing the risk of extinction due to AI should be a global priority, along with risks such as pandemics and nuclear war." Due to these "prospects" regarding new conflicts on September 3-6, 2003, in Rhodes, Greece, the participants of the World Public Forum, in the document "Rhodes Declaration"<sup>29</sup> [44] emphasized the irreversible damage that was inflicted on the Institutions of International Peace and Security, due to unjustified violence in regarding the sovereignty, security and culture of individual nations, the formation of a unilateral peace.

To transform this situation into a just, compassionate and human order will require patience, self-sacrifice and constant activity, without which it is impossible to start solving these global problems. Meeting the basic needs of food, shelter, health, education, work, air and water is a constant priority for all people in all countries at all times.

However, the material, cultural and spiritual consequences of the reckless consumerism paradigm being imposed today and its globalization process are becoming increasingly unacceptable, as they contribute to the accumulation of wealth for a few countries and people and lead to the impoverishment of the majority of the world's population.

Ways must be found to break the vicious connection that provokes internal and external compulsion to consume, acquire and hoard as the driving force of the New Global World Order. Society should glorify spiritual values and achievements. A cultural and spiritual counteroffensive is needed.

In different civilizations, original cultures develop, so the process of forced alignment of different cultures into one civilizational model is detrimental to the evolution of mankind. The dialogue of faiths and cultures can develop freely and creatively only if there is no threat of absorption of

and other nations are developing AI applications for a range of military functions.

<sup>27</sup> CAIS - pronounced 'case' is a San Francisco-based research and field building nonprofit. We believe that artificial intelligence (AI) has the potential to profoundly benefit the world, provided that we can develop and use it safely. However, in contrast to the dramatic progress in AI, many basic problems in AI safety have yet to be solved. Our mission is to reduce societal-scale risks associated with AI by conducting safety research, building the field of AI safety researchers, and advocating for safety standards.

<sup>28</sup> Competition could push nations and corporations to rush AI development, relinquishing control to these systems. Conflicts could spiral out of control with autonomous weapons and AI-enabled cyberwarfare. Corporations will face incentives to automate human labor, potentially leading to mass unemployment and dependence on AI systems. As AI systems proliferate, evolutionary dynamics. (<https://time.com/6283958/darwinian-argument-forworrying-about-ai/>)

<sup>29</sup> "Dialogue of Civilization for a Humane Order" was adopted in which there was a plea to maintain the conversation on the destinies of humankind, as well as an expression of the ambition to make such public dialogue a substantive factor in international cooperation. Since 2003, the Rhodes Forum has been convening every year in the conference; these conferences have been instrumental in the creation of an international network of experts.

cultural diversity into one global standard.

In modern world realities, it becomes obvious that the dialogue of civilization and cultures is becoming a universal hermeneutic genetic code for understanding socio-cultural phenomena and relations that are international, inter-ethnic, inter-confessional and inter-civilizational in nature. The civilizational space in which cultural-cognitive interaction takes place should be understood not as real-time facts, but as an area for the implementation of cultures and civilizations.

In modern society, there is a need for cultural empathy in solving the global problems of our time. The dialogue of civilizations with their cultural, ethnic, and social characteristics has become a vital paradigm of humanity. There is no doubt that inter-civilizational and intercultural dialogue is the most important strategy of modern human civilization. It is the only means that can be directed toward solving the most pressing problems of the modern world for the benefit of future generations.

Researchers at Google recently discovered that artificial intelligence (AI) is capable of designing full-fledged chips, and could potentially do so more efficiently than human experts in the field.

In an article published in the journal *Nature* [45], the Google Research team explains that the process of designing the physical layout of a computer chip is a very laborious and time-consuming, but at the same time very important task. To increase efficiency, the researchers used machine learning technologies.

Google developed a series of algorithms and taught them to think of chip design as a jigsaw puzzle game in which the pieces are the basic components of the future chip. The main goal of the game was to reach a certain threshold of quality and efficiency of the finished scheme. The assessment of these indicators was carried out based on a set of 10 thousand ready-made designs of microcircuits, which the researchers gave to the algorithms in the form of training material before the start of the experiment.

If it can take months to design a chip for a living person, AI algorithms can do this task in just six hours. At the same time, ready-made schemes turned out to be similar, and in some cases even more effective, than those developed by real experts. "Our method allowed us to generate a physical layout of the chip in about six hours. Typically, such tasks require months of work by human experts," the researchers from Google Research noted in their article.

According to the researchers, the method not only reduces development time but also improves the quality of the final result, since the algorithms allow more accurate tasks to be solved by the desired location and combination of the components that make up the microcircuits. In their article, experts from Google Research also noted that the method of using AI has already been put into practice in the development of the latest generation of Google's tensor processor.

The company announced back in 2020 that it was experimenting with AI in the development of its chips. The head of the artificial intelligence department, Jeff Dean, then noted that this approach would reduce the financial costs of

development and at the same time help create more efficient microchip designs.

In continuation of the topic, we can recall that Siemens PLM Software [46] is the world's leading provider of software solutions for product lifecycle management PLM [47] and production MOM [48]. Siemens PLM Software's [49] intelligent portfolio of innovation solutions helps manufacturers optimize their digital manufacturing processes and drive innovation.

Pointed out about innovations by Google Research and Siemens PLM Software highlights the fears of the American "AI Security Center" [50] that many of the main security problems of AI have yet to be solved today<sup>30</sup>.

## 4. Instead of an Epilogue

Out of 100 billion neurons, none are alike. They independently establish connections with each other, exchange information and even talk to each other... Do we believe that we make any decisions ourselves? In fact, the decision to act is made by the brain 0.3 seconds before it is reported to our conscious mind [51]. Based on the above, what will happen if artificial intelligence is used in wars? Humans may no longer be able to control armed robots, and they may begin to make independent decisions.

## 5. Conclusion

The dialogue of cultures is becoming a universal hermeneutic genetic code for understanding sociocultural phenomena and relations that are international, interethnic, interfaith, and inter-civilizational in nature.

The civilizational space, in which cognitive interaction takes place, should be understood not as a real-time reality, but as an area of realization of cultures and civilizations.

In modern world society, there is a need for cultural empathy in solving the global problems of our time.

It is obvious that the dialogue of civilizations with their cultural, ethnic, and social characteristics has become a vital paradigm of humanity.

Inter-civilizational and intercultural dialogue is the most important strategy of modern human civilization. This is the only means that can be directed to solving the most acute problems of the modern world for the benefit of future generations.

I present a project proposal for the international consortium of inter-civilizational and intercultural dialogue.

The functional purpose of such an international consortium is determined by the fundamental problems of the modern world, for example:

- 1) Artificial intelligence is a new threat with the function of weapons of mass destruction.
- 2) Stationary carcinogens as causes of cancer.

---

<sup>30</sup> Our mission is to reduce societal-scale risks associated with AI by conducting safety research, building the field of AI safety researchers, and advocating for safety standards.

- 3) Systemic solutions to overcome protein hunger among the world's population.
- 4) Technological platforms and algorithms for the settlement of interethnic and interstate conflicts.
- 5) Ecological paradigm of the modern world and fractal technological platform.
- 6) New professions for university spaces of the 21st century.
- 7) Natural food compounds for direct-acting drugs against the novel SARS-CoV-2 coronavirus.

Features of intercultural and inter-civilizational dialogue imply, as a result, the search for an answer to some questions. First of all, this is a question of methodological choice for solving the problem of identifying internal and external capabilities and conditions of civilization and cultural communications in almost all spheres of human activity.

Logically interconnected with the above is the question of a subject capable of realizing the possibilities of dialogue in the format of interaction regarding the fundamental problems of our time.

The modern civilizational space, in which the cultural-cognitive interaction takes place, should be understood not as a fact of real time, but as a field of realization of cultures and civilizations. It should be noted that the dialogue of civilizations and cultures is a universal hermeneutic genetic code for the systematic perception of socio-cultural phenomena and interpersonal relations, which in turn is an international, ethnic, inter-religious, and civilizational paradigm. It is obvious that the cognitive dialogue of civilizations with their own cultural, ethnic, and social characteristics has survived as the only vital social system of humanity. In modern society, we can turn the centered teaching of cognitive-cultural dialogue into the most effective mechanism for solving the global problems of our time.

Civilizational and intercultural dialogue appears to be the most important strategy of modern man. This is the only means that can be directed to solving the most pressing problems of the modern world for the benefit of future generations. It should be noted here that the systematic analysis of global crises reveals that the causes of crisis collisions are identified with modern globalization, which generally means the globalization of the world. For mankind, the Lord created a single world and it was a global world and this globality was around the Lord.

Later, however, man directed globalization in another direction - an attempt to unify the world using force without God's commandments. In the history of mankind, many examples of unification by forceful methods without God are known: Alexander the Great (Macedonian), who was obsessed with the ideologeme of the universalism of ancient Greek philosophy (the term "ideologeme" is used in the study of totalitarian discourse).

The totalitarian language is a peculiar system of ideologemes, and it is, for example, "Citizen of Rome," "Soviet man" "or "new Ariel" was a means of forming an ideologized consciousness. The term "mythologeme"

established by K. G. Jung (meaning the creation of new faces using mythological archetypes and the import of old faces and stories into a new context) and tried to unify the world through wars.

The Roman Empire was also able to bring under its protectorate vast territories from the British Isles to the Middle East and from Germany to Africa.

At the dawn of human civilization, the Persian Empire attempted to globalize the world, The Islamic Caliphate expanded from the Atlantic Ocean to China, and so did the Mongol Empire and the Ottoman Empire. Examples of forced globalization in the 20th century are the Soviet communist regime in Russia and Adolf Hitler's attempt to unify the world with the German racial theory.

At the modern stage, new forms and methods of globalization of the world without God have emerged - for example, one of the economic indicators of liberal capitalism is in the form of world money, the expansion of which, for the purpose of their globalization in different countries, gives rise to an integrated world economic system, and this is the main cause of global crises. Everything that goes against the Lord's will always end in systemic crises.

Based on the above, in this article, we present several examples of organizing a platform for the dialogue of civilizations and cultures, we have in mind the presentation of several cases as strategies around which it is possible to design a useful dialogue of representatives of societies of different civilizations and cultures.

One of the missions of the presented paper can be defined as the formation of a systematic topology of intentions to solve the fundamental problems of the modern world.

The audience of the reader of this material: diplomats, sociologists, conflictologists, employees of international organizations, governmental and non-governmental organizations.

## ORCID

0000-0002-3228-1447

## Conflicts of Interest

The author declares no conflicts of interest.

## References

- [1] Ada Lovelace, in full Ada King, countess of Lovelace, original name Augusta Ada Byron, Lady Byron // Available from: <https://www.britannica.com/biography/Ada-Lovelace> [Accessed 22 December 2023].
- [2] Alan Turing // Available from: <https://www.britannica.com/biography/Alan-Turing> [Accessed 22 December 2023].
- [3] Joseph Weizenbaum and the Emergence of Chatbots // Available from: <https://doi.org/10.1093/oso/9780190080365.003.0004> [Accessed 22 December 2023].

- [4] "Deep Blue" defeated Garry Kasparov in chess match // Available from: <https://www.history.com/this-day-in-history/deep-blue-defeat-s-garry-kasparov-in-chess-match> [Accessed 22 December 2023].
- [5] AI Index Report 2023 // Available from: [https://aiindex.stanford.edu/wp-content/uploads/2023/04/HAI\\_AI-Index-Report\\_2023.pdf](https://aiindex.stanford.edu/wp-content/uploads/2023/04/HAI_AI-Index-Report_2023.pdf) [Accessed 22 December 2023].
- [6] International Electronic Commission - IEC // Available from: <https://www.iec.ch/homepage> [Accessed 22 December 2023].
- [7] ISO/IEC TR 24028:2020 standard // Available from: <https://webstore.iec.ch/publication/67138> [Accessed 22 December 2023].
- [8] Chat GPT // Available from: [https://en.wikipedia.org/wiki/ChatGPT#cite\\_note-guardianpos-2](https://en.wikipedia.org/wiki/ChatGPT#cite_note-guardianpos-2) [Accessed 22 December 2023].
- [9] Jan Hatzius, Joseph Briggs, Devesh Kodnani, Giovanni Pierdomenico // Global Economics Analyst the Potentially Large Effects of Artificial Intelligence on Economic Growth (Briggs/Kodnani) // Available from: [https://www.key4biz.it/wp-content/uploads/2023/03/Global-Economics-Analyst\\_-The-Potentially-Large-Effects-of-Artificial-Intelligence-on-Economic-Growth-Briggs\\_Kodnani.pdf](https://www.key4biz.it/wp-content/uploads/2023/03/Global-Economics-Analyst_-The-Potentially-Large-Effects-of-Artificial-Intelligence-on-Economic-Growth-Briggs_Kodnani.pdf) // [https://www.ansa.it/documents/1680080409454\\_ert.pdf](https://www.ansa.it/documents/1680080409454_ert.pdf) [Accessed 22 December 2023].
- [10] How the sudden rise of AI is shaking your white-collar world // Available from: <https://www.afr.com/technology/will-the-sudden-rise-of-useful-ai-shake-our-cosy-white-collar-world-20230126-p5cfnj> [Accessed 22 December 2023].
- [11] The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation // Available from: <https://img1.wsimg.com/blobby/go/3d82daa4-97fe-4096-9c6b-376b92c619de/downloads/MaliciousUseofAI.pdf?ver=1553030594217> [Accessed 22 December 2023].
- [12] The INTERPOL Face Recognition System (IFRS) - global criminal database // Available from: <https://www.interpol.int/How-we-work/Forensics/Facial-Recognition> [Accessed 22 December 2023].
- [13] Artificial Intelligence Index Report 2023 // Stanford University // Available from: [https://aiindex.stanford.edu/wp-content/uploads/2023/04/HAI\\_AI-Index-Report\\_2023.pdf](https://aiindex.stanford.edu/wp-content/uploads/2023/04/HAI_AI-Index-Report_2023.pdf) [Accessed 22 December 2023].
- [14] Zhu, Q., Huang, Y., Zhou, D. *et al.* Automated synthesis of oxygen-producing catalysts from Martian meteorites by a robotic AI chemist. *Nat. Synth* (2023) // Available from: <https://doi.org/10.1038/s44160-023-00424-1> [Accessed 22 December 2023].
- [15] Oren Etzioni. How to know if artificial intelligence is about to destroy civilization // Available from: <https://www.technologyreview.com/2020/02/25/906083/artificial-intelligence-destroy-civilization-canaries-robot-overlord-s-take-over-world-ai/> [Accessed 22 December 2023].
- [16] Artificial Intelligence Index Report 2021 // CHAPTER 7: AI POLICY AND NATIONAL STRATEGIES // Available from: [https://aiindex.stanford.edu/wp-content/uploads/2021/03/2021-AI-Index-Report-\\_Chapter-7.pdf](https://aiindex.stanford.edu/wp-content/uploads/2021/03/2021-AI-Index-Report-_Chapter-7.pdf) [Accessed 22 December 2023].
- [17] Martin V. Butz // Towards Strong AI // Available from: <https://link.springer.com/article/10.1007/s13218-021-00705-x> [Accessed 22 December 2023].
- [18] The law enforcement understand what challenges derivative and generative AI models could pose // Available from: <https://www.europol.europa.eu/cms/sites/default/files/documents/Tech%20Watch%20Flash%20-%20The%20Impact%20of%20Large%20Language%20Models%20on%20Law%20Enforcement.pdf> [Accessed 22 December 2023].
- [19] AI rules as Chat GPT // Available from: <https://www.reuters.com/world/us/senate-leader-schumer-pushes-ai-regulatory-regime-after-china-action-2023-04-13/> [Accessed 22 December 2023].
- [20] Mira Murati // OpenAI's research, product and safety teams // Available from: <https://apnews.com/article/openai-cto-mira-murati-chatgpt-gpt4-dalle-0e701dd406b4a1a2d625e779de0c5164> [Accessed 22 December 2023].
- [21] Geoffrey Hinton is a pioneer of deep learning // MIT Technology Review // Available from: <https://www.technologyreview.com/2023/05/02/1072528/geoffrey-hinton-google-why-scared-ai/> [Accessed 22 December 2023].
- [22] Henry Alfred Kissinger // Available from: <https://www.goodreads.com/book/show/71083945-the-age-of-ai> [Accessed 22 December 2023].
- [23] Henry A. Kissinger. How the Enlightenment Ends Philosophically, intellectually - in every way - human society is unprepared for the rise of artificial intelligence. Available from: <https://www.theatlantic.com/magazine/archive/2018/06/henry-kissinger-ai-could-mean-the-end-of-human-history/559124/> [Accessed 22 December 2023].
- [24] AlphaGo is a narrow AI // Available from: <https://www.deepmind.com/research/highlighted-research/alphago> [Accessed 22 December 2023].
- [25] The idea of applying ethics to artificial intelligence. Available from: <https://www.techopedia.com/definition/32768/deepmind> [Accessed 22 December 2023].
- [26] What Does Alpha Go Mean? Available from: <https://www.techopedia.com/definition/31955/alphago> [Accessed 22 December 2023].
- [27] "Gestalt" // Available from: <https://www.britannica.com/dictionary/gestalt> [Accessed 22 December 2023].
- [28] Natalie Sidamonidze. Artificial intelligence is a challenge and some methodological aspects of its realization. Transactions. Georgian Technical University. AUTOMATED CONTROL SYSTEMS - No 1(28), 2019.
- [29] A robot broke a child's finger at a chess tournament. Available from: <https://www.ixbt.com/news/2022/07/21/robot-slomal-palec-ebenku-na-shahmatnom-turnire-v-moskve.amp.html> [Accessed 22 December 2023].
- [30] Kissinger, H. A. How the Enlightenment Ends / H. A. Kissinger // Journal the Atlantic. 06/01/2018. p. 12-16.

- [31] Go is an abstract strategy board game for two players // Available from: [https://en.wikipedia.org/wiki/Go\\_\(game\)](https://en.wikipedia.org/wiki/Go_(game)) [Accessed 22 December 2023].
- [32] Artificial intelligence chatbot that was originally released by Microsoft Corporation// Available from: [https://en.wikipedia.org/wiki/Tay\\_\(chatbot\)](https://en.wikipedia.org/wiki/Tay_(chatbot)) [Accessed 22 December 2023].
- [33] Alpha Zero is a computer program developed by artificial intelligence research company DeepMind to master the games of chess, shogi, and go. Available from: <https://en.wikipedia.org/wiki/AlphaZero> [Accessed 22 December 2023].
- [34] Artificial Intelligence (AI) technology // Available from: <https://futureoflife.org/open-letter/open-letter-autonomous-weapons-ai-robotics/> [Accessed 22 December 2023].
- [35] Irving John Good // Available from: [https://en.wikipedia.org/wiki/I.\\_J.\\_Good](https://en.wikipedia.org/wiki/I._J._Good) [Accessed 22 December 2023].
- [36] Large-scale risks and towards benefiting life// Available from: <https://futureoflife.org/about-us/> [Accessed 22 December 2023].
- [37] the Future of AI // Available from: <https://futureoflife.org/event/ai-safety-conference-in-puerto-rico/> [Accessed 22 December 2023].
- [38] Stuart Russell, Daniel Dewey, Max Tegmark // Research Priorities for Robust and Beneficial Artificial Intelligence. // Available from: [https://futureoflife.org/data/documents/research\\_priorities.pdf?x89399](https://futureoflife.org/data/documents/research_priorities.pdf?x89399) [Accessed 22 December 2023].
- [39] Research Priorities for Robust and Beneficial Artificial Intelligence // Available from: <https://futureoflife.org/open-letter/ai-open-letter/> [Accessed 22 December 2023].
- [40] National Security and Artificial Intelligence // Available from: <https://novapublishers.com/shop/national-security-and-artificial-intelligence/> [Accessed 22 December 2023].
- [41] The Center for AI Safety // Available from: <https://www.safe.ai> [Accessed 22 December 2023].
- [42] four key categories of catastrophic AI risk// Available from: <https://www.safe.ai/ai-risk> [Accessed 22 December 2023].
- [43] AI race - safety regulations, international coordination, and public control of general-purpose AI's // Available from: <https://www.safe.ai/ai-risk> [Accessed 22 December 2023]
- [44] The Rhodes Declaration // Available from: [https://en.wikipedia.org/wiki/Dialogue\\_of\\_Civilizations](https://en.wikipedia.org/wiki/Dialogue_of_Civilizations) [Accessed 22 December 2023]
- [45] A graph placement methodology for fast chip design// Available from: <https://www.nature.com/articles/s41586-022-04657-6> [Accessed 22 December 2023].
- [46] Siemens Digital Industries Software portfolio // Available from: <https://www.plm.automation.siemens.com/global/en/products/> [Accessed 22 December 2023].
- [47] Product Lifecycle Management – PLM // Available from: <https://www.plm.automation.siemens.com/global/en/industries/automotive-transportation/automotive-suppliers/electronics-manufacturing-planning.html> [Accessed 22 December 2023].
- [48] Manufacturing Operations Management – MOM // Available from: <https://www.plm.automation.siemens.com/global/en/resource/manufacturing-operations-management-with-opcenter-execution-discrete/107834> [Accessed 22 December 2023].
- [49] CAD software for mechanical design // Available from: <https://www.plm.automation.siemens.com/global/en/our-story/offices.html> [Accessed 22 December 2023].
- [50] Basic problems in AI safety have yet to be solved. Available from: <https://www.safe.ai> [Accessed 22 December 2023]
- [51] Ramaz Khetsuriani: The brain has such abilities that should not be the result of evolution! (In Georgian) // Available from: <https://iberiana.wordpress.com/health/khetsuriani/> [Accessed 22 December 2023].