

## „QUO VADIS AI?“

Emil Asenov<sup>1</sup>

e-mail: [emilassenov@unwe.bg](mailto:emilassenov@unwe.bg)

### Abstract

*This article explores the challenges that arose in the context of changes in the corporate governance of OpenAI, the company behind the development of ChatGPT. The internal conflict within the corporation clearly reveals a confrontation between two opposing ideological trends. On one hand, there are representatives of “effective altruism” who believe that uncontrolled development of artificial intelligence is linked to the emergence of existential threats to the future of human civilization. On the other hand, a completely different position is taken by the “effective accelerationists” who argue that consciously and purposefully stimulating technological development is necessary, ignoring possible negative effects. This conflict leads some analysts to claim that the world is witnessing the “first battle in the war for artificial intelligence.”*

*With the help of artificial intelligence, humans will attempt to build a new and more perfect world. However, the world of the future will no longer belong solely to humans. It will be a shared world with artificial intelligence, where the human and technological elements will intertwine. Technological development and the strong presence of artificial intelligence will dramatically transform the way people live.*

*The major unknown at this moment is to what extent artificial intelligence can be one of the most powerful tools for technological development, or conversely, how it can be a source of new, previously unknown risks that humanity is unable to manage.*

**Keywords:** artificial intelligence, OpenAI, ChatGPT, effective altruism, effective accelerationism

**JEL:** O32, O33, K24

In the vision regarding the future development of artificial intelligence shared by the specialists developing it, corporate leaders, and major investors, there is a clear confrontation between two, until now unknown, ideological trends. The likely basis of their conflict lies at the heart of the dramatic events that unfolded around OpenAI, the company behind the development of ChatGPT.

Adherents of one of these ideologies believe that the rapid development of artificial intelligence could quickly lead to the downfall of humanity. Another faction, rather, insists on the unrestricted development of artificial intelligence,

---

<sup>1</sup> Prof., PhD, Department of Financial Control, Faculty of Finance and Accounting, University of National and World Economy, Sofia

even if it is associated with the emergence of existential risks for the future of contemporary civilization. It turns out that both groups have almost irreconcilable and influential supporters in technological circles. The events at OpenAI in late November 2023 provided grounds for some analysts to claim that the world has witnessed the “first battle in the war for artificial intelligence” (Nitish, 2023).

The aim of this article is to analyze the differences underlying the ideological confrontation regarding the future technological development of the world. In this context, a review and analysis of the main ideological perspectives within the technological community will be conducted.

The main theses of the study are as follows:

1. For the first time in history, there is a sharp ideological confrontation within the field of a new information technology. It is somewhat paradoxical that this is not so much triggered by a crisis or unresolved problems, but rather by the rapid and perhaps unexpected success in the development of the technology. Although at first glance the opposition appears to be within a specific corporation, the outlined lines of division far exceed its boundaries and spread within the global technological community.
2. The corporate conflict in OpenAI brings to the forefront an important existential question for the future of humanity. To what extent will priority be given in the future to new technological innovations that could lead to radical social changes, or will the familiar approach of creating regulatory frameworks to manage potential risks be adopted? The actions of many countries and international organizations currently indicate that active steps are being taken to regulate the new and extensive domain related to artificial intelligence.

### **Brief Chronology of Events in the Leadership of OpenAI**

The news of the unexpected dismissal on Friday, November 17, 2023, of a co-founder of OpenAI was surprising and logically shook Silicon Valley. The Board of Directors of OpenAI removed Sam Altman, who was the CEO of the company and one of its co-founders. The decision to dismiss him was made very quickly and without prior warning. It turned out that neither Altman, nor the major investors, nor even Microsoft's leadership, which owns 49% of OpenAI's shares, had any information about the impending dismissal until the last moment.

The unusual, official announcement of Altman's dismissal also raised many questions. In the announcement, the Board of Directors of OpenAI directly accused the former executive director of lying. Such aggressive formulations are rare, even when high-level managers are dismissed for crimes they have committed. The announcement did not specify any facts that would clearly indicate what Altman was specifically accused of lying about. Altman's duties

were temporarily assigned to Mira Murati, who was previously the technical director of OpenAI.

Over the weekend, the Board of Directors of OpenAI attempted to negotiate with a series of investors. As a result, on Sunday, November 19, 2023, the management of the company was once again changed, and the head of the video service Twitch, Emmett Shear, was appointed as the interim director.

Meanwhile, Altman, his close associate Greg Brockman, and several leading experts announced that they were moving to work at Microsoft. The next day, on Monday, November 20, 2023, another unexpected event occurred: 700 out of a total of 770 employees issued an ultimatum to the Board of Directors, demanding the return of Altman. They stated that they would leave the company if their request was not fulfilled. Among the signatories of the open letter were Mira Murati and even the scientific director of OpenAI, Ilya Sutskever, one of the main actors in these events, and according to unofficial information, the initiator of Altman's dismissal.

On Tuesday evening, November 21, 2023, Altman agreed to return to OpenAI. On the same day, the Board of Directors of OpenAI announced their collective resignation.

**The leadership of OpenAI was deeply concerned about the speed at which the company's developments related to artificial intelligence were progressing.**

OpenAI is considered one of the most important players in the field of artificial intelligence. Among its projects are the large language model GPT, the ChatGPT service that allows users to interact with GPT in a chat mode, and the DALL·E service, which generates images based on textual descriptions. However, the company has always had a more ambitious goal - to create a universal artificial intelligence (UAI) that surpasses human intelligence. A few years ago, universal artificial intelligence seemed like a distant and almost unattainable goal. At that time, the founders of OpenAI were willing to admit that their research might not be successful and could be a source of financial losses. For this reason, the company was established as a non-profit research laboratory.

“Serious problems, however, arose as a consequence of the fact that OpenAI's technologies started developing too rapidly. After the release of ChatGPT, a clear path to a sharp increase in revenue and profits was discovered. It became difficult to explain why the company should continue to be a non-profit research laboratory when many interested parties had already emerged, demanding quick financial results”.

The unexpected success of ChatGPT unequivocally demonstrated that the initial forecasts were entirely wrong. According to information from internal sourc-

es, the pace of changes has always been a cause for concern, even for the creator of GPT, Ilya Sutskever. He genuinely believed in the need to create superhuman intelligence and became a kind of corporate and spiritual leader (Hao, Warzel, 2023), guiding the company toward this goal.

The rapid commercial success of OpenAI's products, a goal pursued by Altman's leadership, did not allow for the necessary pause in the development process that influential groups in society were urging. Many analysts, businessmen, and public figures stated that it is critically important to urgently analyze and contemplate the ultimate impact that artificial intelligence will have on the fate of humanity. According to some sources, the catalyst for the dramatic sequence of events in the company was the developed algorithm Q\*.

According to some information, the Q\* algorithm turned out to be significantly more effective than expected, significantly increasing the likelihood of the rapid emergence of universal artificial intelligence (Tong, Dastin, Hu, 2023). In November, the developers of Q\* warned the Board of Directors of OpenAI about the potential threat posed by this algorithm. It is important to note that this warning came directly from the developers of the algorithm and not from the CEO, Altman.

A few days later, Altman's unexpected dismissal without warning became a reality. Opponents of the accelerated development of artificial intelligence, including Sutskever, likely believed that artificial intelligence surpassing human capabilities had reached a point of no return, posing a real existential threat. Initially, OpenAI had hoped to manage this danger through regulation (alignment) – built-in safety mechanisms preventing the intentions of artificial intelligence from deviating from human motives and goals in behavior. However, the rapid development of artificial intelligence increasingly showed that no safety mechanisms would be able to contain it within acceptable and safe bounds. Such a course of events clearly led to the emergence of the greatest danger, according to Sutskever.

During the summer, he, along with over a hundred prominent researchers in the field of artificial intelligence, businessmen, investors, and public figures, including Elon Musk and Steve Wozniak, published an open letter<sup>2</sup>. In it, they raised the question of the need for a responsible approach to the risks associated with artificial intelligence. According to the authors of the letter, we should treat this new danger with the same seriousness as we approach nuclear or biological threats. Specifically, it was emphasized that “powerful artificial intelligence systems should only be developed when we are certain that the effects of their deployment will be positive, and the risks are manageable”.

---

<sup>2</sup> Mitigating the risk of extinction from AI should be a global priority alongside other societal-scale risks such as pandemics and nuclear war (CAIS, n.d.).

Regarding his stance on issues related to the regulation of artificial intelligence, there is information that Sutskever ritually burned a doll figure representing “unregulated artificial intelligence” during a meeting with the leadership of OpenAI. This doll figure was commissioned from a local artist (Hao, Warzel, 2023).

As the most radical yet influential advocate of the ideological movement opposing the rapid development and implementation of artificial intelligence, Eliezer Yudkowsky, the founder of the Machine Intelligence Research Institute, is often mentioned. Yudkowsky insists not only on the immediate cessation of OpenAI but also on any research in the field of artificial intelligence. He calls on public authorities not to spare any effort in pursuing such goals (Yudkowsky, 2023). For example, if graphics cards are necessary for training artificial intelligence, he suggests strict penalties for merely owning graphics cards. According to Yudkowsky, the United States should be prepared, if other countries violate the prohibition standards, to bombard their technological centers. He believes that if necessary, even nuclear weapons could be used. He argues that a nuclear war can be survived, but one involving artificial intelligence cannot. While Yudkowsky’s reasoning is highly radical, it has garnered significant public attention in Silicon Valley.

People with similar views are encountered in many places, including OpenAI. In particular, Yudkowsky has personal communication with Emet Shir, appointed as the temporary director of OpenAI for a short period, who openly identifies as a “thoughtful” person. In principle, he shares Yudkowsky’s apocalyptic views, although his ideas for responding to issues related to artificial intelligence are more moderate. He does not propose a complete halt to research but rather advocates significantly slowing down its pace. “If the current pace is 10, and the pause is zero pace, according to him, we should aim for a pace of 1-2”, he wrote on Twitter.

Several other members of OpenAI’s board of directors have been identified as supporters of the ideology broadly termed “effective altruism”. According to this ideology, societal interests are always more important than individual interests. In this context, the main goal shared by advocates of “effective altruism” is to prevent threats arising from the uncontrolled development of artificial intelligence.

Supporters of this ideology on the OpenAI board of directors also include Helen Toner from the Center for Security and Emerging Technology at Georgetown University and the Oxford Center for Governance of Artificial Intelligence, representing the Effective Altruism Center, and Tasha McCauley, who works at the RAND Corporation research center.

After a new Board of Directors was elected at OpenAI at the end of November 2023, Microsoft announced that it would have a non-voting seat on the board. This move was logical considering that the future development of OpenAI would continue to closely involve collaboration with Microsoft. The software giant invested \$13 billion in OpenAI and integrated its artificial intelligence models into its software products.

Mira Murati, who was the Chief Technology Officer at OpenAI and briefly appointed as the interim CEO, returned to her previous position. Greg Brockman, on the other hand, resumed his role as the President of OpenAI.

Former CEO of Salesforce, Bret Taylor, who will lead the new board, according to a statement published on OpenAI's website, mentioned that he would focus on "strengthening OpenAI's corporate governance." In a subsequent post on X, Taylor commented that he would leave the board once it is filled with people, and the company stabilizes (Kirilova, 2023).

### **Regarding the ideas, principles, and goals of "effective altruists"**

The significant presence of representatives of "effective altruism" in the governing body of OpenAI, which replaced Altman, raises questions about the goals, priorities, and characteristics of this ideological movement. In essence, it is a philosophical and social movement that advocates for the "use of evidence and reason to find approaches to benefit others as much as possible and take corresponding actions" (Hao, Warzel, 2023).

Effective altruists identify the most critical philosophical questions concerning the most effective ways to benefit others. Therefore, they often engage in activities related to achieving ideal goals. The goals and approaches of effective altruism strongly influence the planning and implementation of scientific projects, public corporate activities, and political initiatives. Achieving the ultimate goals of effective altruism involves ensuring people's safety, improving the quality of life, or maximizing benefits for individuals (MacAskill, 2017).

The philosophical principles of effective altruism include impartiality, cause neutrality, cost-effectiveness, and comparative analysis ("opposite reasoning").

In the ideological landscape of effective altruism, there are also other similar ideological currents referred to by terms such as "AI doomers" or "decelerationists".

As "doomers", individuals are characterized by extreme pessimism, even fatalism, regarding the prospects of finding solutions to global problems related to overpopulation, climate issues, environmental pollution, nuclear weapons, and the rapid development of artificial intelligence. Some doomers argue that there is a real possibility that these problems could lead to the extinction of humanity (BBC News, 2023).

Authors such as Paul R. Ehrlich, Guy McPherson, and Michael Rupert link doomism to Malthusianism – an economic philosophy that posits human consumption will surpass the Earth’s available natural resources, leading to societal collapse, social disorder, or a decrease in population (Grist, 2020; Holmgren, 2009).

On the other hand, “decelerationists” express serious concerns about the fact that technologies are advancing at a pace that cannot be controlled. Representatives of this ideological movement believe that it is crucial to carefully monitor the dynamics of the process regarding the development, implementation, and use of AI.

### **Ideas, Values, and Goals of “Effective Accelerationists”**

In light of the events at OpenAI, the actions of supporters from the camp of “effective altruists” faced sharp criticism. Representatives of the investment community (Paresh, 2023). stated that concerns about the management of OpenAI’s Board of Directors had always existed, but they had never entertained the hypothesis that these individuals would become activists ready to disrupt everything for the sake of an abstract idea. According to other figures, such as the renowned investor Vinod Khosla, (McMillan, 2023) the “religion of effective altruism” followed by the members of OpenAI’s board and its misapplication created a barrier to the world’s progress toward the incredible benefits of artificial intelligence”.

Against the “effective altruists” and their variations, the “doomers”, and “decelerationists”, a powerful ideological and radical faction emerged, self-identifying as “effective accelerationists” or “e/accs”. They perceive themselves as representatives of technocrats and oppose the actions of the “altruists””, who, according to them, interfere with critical technological matters and promote erroneous ideas about ethics and societal well-being.

Speculations surrounding Altman’s dismissal shed light on the deep, practically opposite differences in views regarding the development of artificial intelligence within the technological community. Essentially, it is divided into opposing camps: representatives of effective altruism and effective accelerationism.

Effective accelerationism is an ideology and social movement built on the foundation of the theory of accelerationism, whose intellectual roots can be traced to the works of the French post-structuralists Gilles Deleuze and Félix Guattari. Supporters of this concept are usually positioned on the far right of the conditional political coordinates system. The theory of accelerationism gained popularity actively in the 1990s through the efforts of the British philosopher Nick Land, who chose the term “neoreaction” to describe his views.

Supporters of accelerationism believe that society should progress through the application of new aggressive technological innovations, which should not be subject to additional regulatory frameworks due to concerns about their potential negative effects. The main argument is that new technologies can lead to such



radical social changes, and the long-term benefits could be of such magnitude that potential risks can simply be neglected.

In the recently published programmatic text of accelerationists titled the “Techno-Optimist Manifesto” (Andreessen, 2023). They state that they believe in acceleration – the conscious and targeted stimulation of technological development to ensure the application of the Law of Accelerating Returns, allowing the spiral of growth of techno-capital to continue indefinitely.

The manifesto contains a specific statement that allows positioning this movement within the field of the contemporary ideological spectrum. Specifically, it declares the following: “We have enemies. Our enemies are not bad people, but rather bad ideas. Our modern society for six decades has been subjected to mass demoralization – against technology and against life – under various names: ‘existential risk’, ‘sustainability’, ‘ESG’, ‘sustainable development goals’, ‘social responsibility, stakeholder capitalism’, ‘precautionary principle’, ‘trust and safety’, ‘technoethics’, ‘risk management’, ‘anti-growth’, ‘limits’”.

In response to the assertion by AI doomers that humanity is digging its own grave with the development of new AI technologies, the supporters of “e/acc” provide the following answer: “... since one cannot make much money from doomerism (except by writing good fiction), this is more of an academic position than the stance that technology companies will take. As far as I can judge, ‘e/acc’ is winning in the current discussions, mainly due to rapidly advancing AI technologies and the fact that a significant part of the economy wants to save money, make money, and win more” (Wilhelm, 2023).

This direction has a large number of followers, including former Google CEO Eric Schmidt and renowned investor Marc Andreessen, who developed the first popular web browser years ago. After Altman’s dismissal, Andreessen launched sharp accusations on Twitter against the “effective altruists”. With a significant amount of candor, he pointed out that contrary to the altruists, the “effective accelerationists” believe that the development of artificial intelligence should not only not be slowed down but rather accelerated – even if it leads to the end of the world. As explained by one of the popular ideologists of “e/acc” using the pseudonym “*Beff Jezos*” (Jezos, 2022). Accelerationists do not view the “biological substrate for intelligence and life” particularly favorably (Lenta.ru, 2023).

### **Economic Perspectives on the Development of Artificial Intelligence**

The outcome of the events in the leadership of OpenAI at first glance demonstrates the victory and consolidation of the technocratic approach of the accelerationists, linked to the expansion of the application of artificial intelligence, particularly ChatGPT. There can be little doubt that the economic sector associated with artificial intelligence will undergo rapid development. It might even



be argued that it will significantly dictate the pace of development for the entire economy over the next decade. According to expert forecasts, the global market for artificial intelligence solutions is projected to grow from \$196.6 billion to \$1.8 trillion USD in the period from 2023 to 2030, with an average annual growth rate of 37% (Bloomberg, 2023).

According to McKinsey's estimates, the overall economic potential of implementing artificial intelligence in various industries ranges between \$17.1 trillion and \$25.6 trillion USD. This growth will be driven by increased business productivity through the automation of work processes, enhanced labor productivity, and increased demand for AI-powered products due to the improved quality of these products (McKinsey, 2023).

This effect will be made possible, including by reducing the costs of implementing artificial intelligence models in organizational activities and by increasing access to data for training such models.

The development of generative artificial intelligence will be a leading technological trend associated with the use of artificial intelligence. This includes the development and widespread deployment of generative artificial intelligence, including large language models. Generative artificial intelligence is used to create content such as text, audio, and images.

According to experts, in the coming years, generative artificial intelligence will become one of the most dynamically evolving branches of artificial intelligence. More and more companies are recognizing the competitive advantages of AI for business and making the adoption of this technology a priority (Bakerdzhieva, 2023). Bloomberg Intelligence estimates indicate that the global market for generative AI-based solutions will grow more than 13 times, from \$67 billion to \$897 billion USD, in the period from 2023 to 2030. The growth of public interest in the field of artificial intelligence in recent years is significantly connected to the development of generative AI models (Bloomberg, 2023).

Thanks to the rapid growth in the productivity of generative AI models, they have broad potential for applications in various fields. According to McKinsey's estimates, the economic impact of implementing generative AI products across different sectors of the global economy will reach between \$2.6 trillion and \$4.4 trillion USD annually. The use of generative artificial intelligence in marketing, customer services, and software development constitutes 75% of this growth.

The outlined forecasts provide sufficient grounds to assert that the ideas of accelerated technological development advocated by the effective accelerationists will likely be further solidified in practice. In the coming years, a profound transformation is anticipated not only in a wide range of economic sectors but also in society as a whole.

At the same time, it's crucial not to forget the broader societal context in which the corporate events of OpenAI unfolded. In early November 2023, at the initiative of the United Kingdom, a two-day high-level meeting on artificial intelligence safety took place in Bletchley Park near London. The forum brought together governments, leading companies developing AI systems, civil society organizations, and numerous analysts and experts. Nationally, the majority of technologically advanced countries<sup>3</sup> participated in the forum and endorsed a joint declaration (The Bletchley Declaration, 2023).

The declaration emphatically states the establishment of shared consensus and responsibility regarding the risks, opportunities, and future processes for international collaboration on the safety and research of artificial intelligence. The declaration highlights a common position that all participants play a role in ensuring the safety of artificial intelligence: nations, international forums, and other initiatives, companies, civil society, and academic institutions.

The necessity for issues related to safety to be addressed throughout the entire life cycle of artificial intelligence is confirmed in this document. According to the declaration, those countries and organizations developing highly powerful and potentially harmful AI systems bear a particular responsibility for ensuring the safety of these artificial intelligences. In this regard, efforts of participants striving to provide appropriate transparency and responsibility regarding their plans for measuring, monitoring, and limiting potentially harmful capabilities and their associated effects are supported.

The text of the Bletchley Park declaration shows that the global political elite is inclined to adhere to the leading ideas and principles of "effective altruism" to prevent potential threats arising from uncontrolled development of artificial intelligence.

### **The First Battle for Artificial Intelligence as the Beginning of Defining the Meaning of Technological Development**

The ideological clash within OpenAI should not be interpreted unilaterally. It clearly outlines a new line of conceptual division and even open opposition within technological circles, particularly among professionals directly engaged in the development of artificial intelligence. However, this ideological opposition extends far beyond and involves broader societal circles and stakeholders, including politicians, corporate figures, financiers, the non-profit sector, the academic community, and more.

---

<sup>3</sup> Among them are Australia, Brazil, Canada, China, the European Union, France, Germany, India, Ireland, Israel, Italy, Japan, the Netherlands, Republic of Korea, Singapore, Spain, Switzerland, United Kingdom, and the United States. /a.n/.

The ideas and priorities of “effective altruism” find broad support among representatives in the information technology sector, political leaders, non-governmental organizations, and scientific communities.

Conversely, the leading ideas of “effective accelerationists” shape the positions of the majority of researchers and developers in artificial intelligence systems, the corporate elite, and the financial and investor community.

The given factual situation leads to the following conclusions:

1. The sharp ideological confrontation related to the development of artificial intelligence is the first to emerge within the field of a specific information technology.
2. The peculiarity of this case is that it is not so much triggered by an existing crisis or unresolved problems, but rather by the rapid and perhaps unexpected success.
3. Although the level of opposition at first glance appears to be within a single corporation, the outlined lines of division far exceed its boundaries and spread within the global technological community.
4. The brief unsuccessful attempt to rotate the governing body of OpenAI has brought to the forefront an important existential question for the future of humanity. To what extent will priority be given in the future to new technological innovations that could lead to radical social changes, or will the familiar approach of creating regulatory frameworks to manage potential risks be adopted? The main argument is that new technologies could lead to such radical social changes, and the long-term benefits could be so substantial that potential risks might simply be disregarded.

Until now, the concept of “technological revolution” had a neutral connotation. Now it has its own value and ideological equivalent.

Undoubtedly, the conflict between these two groups, for various reasons and aspects related to artificial intelligence activities, will persist. New ideological factions and ethical phenomena may emerge. One of the main themes of the debates will concern the approach to preventing potential risks and threats arising from the development of artificial intelligence. Should it be the familiar approach of building continuously expanding regulations for artificial intelligence processes and systems<sup>4</sup>, utilizing ethical principles and criteria from the arsenal of effective altruism? Or, conversely, should it be the approach proposed by accelerationists, tackling existing technological problems with more technology and the application of new technological solutions?

---

<sup>4</sup> According to the Stanford University „AI Index Report“ for 2023, the annual number of AI-related laws passed in the 127 surveyed countries has surged from one law enacted in 2016 to 37 laws enacted in 2022 (Vincent, 2023; Maslej et al., 2023).

At the moment, it's not possible to say whether a middle ground can be found that integrates both approaches and provides a universally acceptable solution for managing the risks associated with artificial intelligence.

Apart from the current and future ideological battles, there lies a fierce competitive struggle to achieve technological supremacy among the three most significant global technology hubs. It's difficult to assess how the rivalry between the USA, China, and Europe will unfold and what the outcome will be regarding the attainment of global technological leadership in this field.

The major uncertainty at this moment is the extent to which universal artificial intelligence can become the new technological demiurge of the world we will inhabit in the new millennium. Conversely, it remains unclear how possible it is for universal artificial intelligence to emerge as a source of new, previously unknown problems for which humanity lacks the necessary tools and capacity to solve.

This question is of paramount importance, and according to Henry Kissinger, who believes that "perhaps the greatest challenge is philosophical. The more artificial intelligence enters conceptual and analytical aspects of activities inaccessible to the human mind, the less understandable its processes, impact, and results of its activities will be for humans. If it is concluded that the assistance of artificial intelligence is needed..., the delegation of critical decisions and functions to machines may become inevitable" (Forbes.ru, 2022).

The corporate conflict within OpenAI raises, in my opinion, crucial and fundamental questions about the future of the world in which we will live. "The first battle for artificial intelligence" may turn out to be the beginning of a new ideological divide regarding the definition of priorities and goals of contemporary civilization in the face of accelerating technological development. Humans understand and shape the world and themselves in accordance with the meanings and values they carry. With the help of artificial intelligence, they will attempt to build a new and more perfect world. However, the world of the future will no longer solely belong to humans. It will be a shared world with artificial intelligence, where the human and technological elements will intertwine. Technological development and the strong presence of artificial intelligence will dramatically transform the way people live. Once again, they will be forced to seek an answer to the question of how much their new existence and the meaning of their existence align with their conception of a happy life.

## References

- Kirilova, E. (2023). Microsoft получава място без право на глас в борда на OpenAI, Инвестор, (Kirilova, E., 2023, Microsoft poluchava myasto bez pravo nag las v borda na OpenAI, available at: <https://www.investor.bg/a/566->

- novini-i-analizi/385363-microsoft-poluchava-myasto-bez-pravo-na-glas-v-borda-na-openai
- Forbes.ru. (2022). ИИ массового поражения: Генри Киссинджер и Эрик Шмидт об угрозах кибервойны, available at: <https://www.forbes.ru/tekhnologii/468437-ii-massovogo-porazenia-genri-kissindzer-i-erik-smidt-ob-ugrozah-kibervojny>
- Lenta.ru. (2023). Создатель ChatGPT мог стать жертвой сектантов, ждущих конца света из-за ИИ. Почему им верят миллиардеры и ученые?, 24 ноября, available at: <https://lenta.ru/news/2023/11/24/openai/>
- Andreessen, M. (2023). Techno-Optimist Manifesto, available at: <https://a16z.com/the-techno-optimist-manifesto/>
- Bakerdzhieva, D. (2023). How AI Changes The World, Knowledge Society And 21st Century Humanism, pp. 268-274, available at: <https://buditeli.unibit.bg/images/proceedings/2023/d-bakyrdzieva.pdf>
- BBC News. (2014). Stephen Hawking warns artificial intelligence could end mankind, Retrieved 20 October 2023
- BizNews.com. (2023). Humans will be extinct by 2026 – doom-and-gloom prophet Professor Guy McPherson on abrupt climate change, 20 January, Retrieved 20 October 2023.
- Bloomberg. (2023). Artificial Intelligence Market to Hit \$ 1,811.75 Billion by 2030: Grand View Research.
- Bloomberg. (2023). Generative AI to Become a \$ 1.3 Trillion Market by 2032.
- CAIS. (n.d.). Statement on AI Risk, AI experts and public figures express their concern about AI risk, Mitigating the risk of extinction from AI should be a global priority alongside other societal-scale risks such as pandemics and nuclear war, available at: <https://www.safe.ai/statement-on-ai-risk>
- Department for Science, Innovation & Technology. (2023). The Bletchley Declaration by Countries Attending the AI Safety Summit, 1-2 November, available at: <https://www.gov.uk/government/publications/ai-safety-summit-2023-the-bletchley-declaration/the-bletchley-declaration-by-countries-attending-the-ai-safety-summit-1-2-november-2023>
- Grist. (2020). Only 2020 could bring us words like these, Grist, available at: <https://grist.org/climate/2020-climate-words-of-the-year-doomer-net-zero-anthropause/>
- Jezos, B. (2022). Notes on e/acc principles and tenets, A physics-first view of the principles underlying effective accelerationis, available at: <https://beff.substack.com/p/notes-on-eacc-principles-and-tenets>
- Hao, K. and Warzel, C. (2023). Inside the Chaos at OpenAI, Sam Altman's weekend of shock and drama began a year ago, with the release of Chat-

- GPT, Theatlantic, available at: <https://www.theatlantic.com/technology/archive/2023/11/sam-altman-open-ai-chatgpt-chaos/676050/>
- Holmgren, D. (2009). Future scenarios: how communities can adapt to peak oil and climate change, White River Junction, Vermont, OCLC 1021809104.
- MacAskill, W. (2017). Effective altruism: introduction, *Essays in Philosophy* 18 (1), DOI: 10.7710/1526-0569.1580
- MacAskill, W. (2016). *Doing Good Better: How Effective Altruism Can Help You Help Others, Do Work that Matters, and Make Smarter Choices about Giving Back*, New York, Avery, [2015] pp. 179-195, OCLC 932001639.
- Maslej, N., et al. (2023). Artificial Intelligence Index Report 2023/Chapter 6: Policy and Governance (PDF). AI Index. 2023. Institute for Human-Centered AI, Stanford University, Retrieved 19 June 2023.
- McMillan, R. (2023). How a Fervent Belief Split Silicon Valley – and Fueled the Blowup at OpenAI, Sam Altman’s firing showed the influence of effective altruism and its view that AI development must slow down; his return marked its limits, *WSJ*, available at: <https://www.wsj.com/tech/ai/openai-blowup-effective-altruism-disaster-f46a55e8>
- McKinsey. (2023). The economic potential of generative AI: The next productivity frontier.
- Nitish, P. (2023). What the Heck Just Happened at OpenAI?? The sacking of Sam Altman is the first skirmish in a war over A.I, available at: <https://slate.com/technology/2023/11/openai-sam-altman-ai-microsoft-eacc-effective-altruism.html>
- Paresh, D. (2023). How OpenAI’s Bizarre Structure Gave 4 People the Power to Fire Sam Altman, *Wired*, Nov 19, available at: <https://www.wired.com/story/openai-bizarre-structure-4-people-the-power-to-fire-sam-altman/>
- Tong, A., Dastin, J. and Hu, K. (2023). OpenAI researchers warned board of AI breakthrough ahead of CEO ouster, sources say, *Reuters*, available at: <https://www.reuters.com/technology/sam-altmans-ouster-openai-was-precipitated-by-letter-board-about-ai-breakthrough-2023-11-22/>
- Yudkowsky, E. (2023). Pausing AI Developments Isn’t Enough. We Need to Shut it All Down, March 29, *Time*, available at: <https://time.com/6266923/ai-eliezer-yudkowsky-open-letter-not-enough/>
- Vincent, J. (2023). AI is entering an era of corporate control, *The Verge*, Retrieved 19 June 2023.
- Wilhelm, A. (2023). <https://techcrunch.com/author/alex-wilhelm/page/4/>