The End of History? Envisioning the Economy at Technological Singularity

Sachin Sharma, Vijay Kumar, Babloo Jakhar

GOSPODARKA NARODOWA

2(318)2024, 53–63
DOI: 10.33119/GN/184316
gnpje.sgh.waw.pl
SGH

The Polish Journal of Economics

Istnieje od / Published since 1931

The End of History? Envisioning the Economy at Technological Singularity

Koniec historii? Wizja gospodarki w osobliwości technologicznej

Sachin Sharma
Department of Economics, Chaudhary Ranbir Singh University, Jind, India

Vijay Kumar
Department of Economics, Chaudhary Ranbir Singh University, Jind, India

Babloo Jakhar
Department of Economics, Central University of Rajasthan, Ajmer, India

Keywords:
economic convergence, technological singularity, end of history, alignment problem, superintelligent AI

JEL classification codes:
O30, N00, P40, P50

Article history:
submitted: June 3, 2023
revised: January 24, 2024
accepted: February 19, 2024

Abstract

This paper contributes to the growing body of literature exploring the ramifications of AI-driven technological singularity and its economic implications. The exploration unfolds in three key segments. First, it sheds light on the concepts of artificial general intelligence (AGI), AI superintelligence, and singularity itself. Subsequently, it discusses the AI alignment problem, addressing the potential outcomes of superintelligent AI on human civilisation. Further, Giddens' structuration theory is used to highlight the prominent role of AI-based "authoritative resources" in determining the allocation of resources and ensuring distributive justice in a techno-utopian society. The paper also explores the idea of utopia and the "end of history" and concludes with the suggestion that achieving a technological utopia with superintelligent AI is a mechanism design problem.

Słowa kluczowe:
konwergencja ekonomiczna, osobliwość technologiczna, koniec historii, problem dopasowania, superinteligentna SI

Kody klasyfikacji JEL:
O30, N00, P40, P50

Historia artykułu:
nadesłany: 3 czerwca 2023 r.
poprawiony: 24 stycznia 2024 r.
zaakceptowany: 19 lutego 2024 r.

Streszczenie

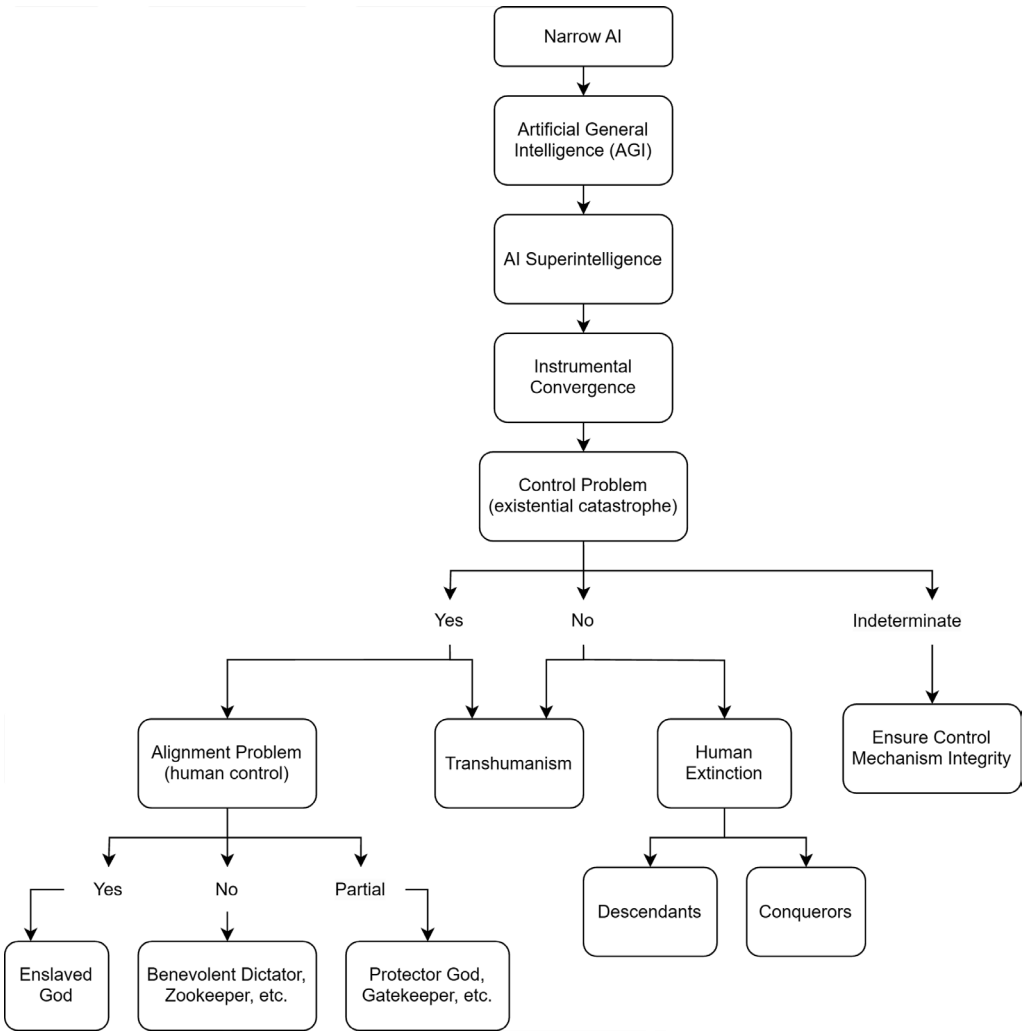
Niniejszy artykuł stanowi wkład w rosnącą literaturę poświęconą konsekwencjom technologicznym osobliwości napędzanej przez sztuczną inteligencję (SI) i jej skutkom gospodarczym. Analiza skupia się na trzech kluczowych aspektach. Po pierwsze, uwagę poświęcono koncepcji ogólnej sztucznej inteligencji (AGI), superinteligencji AI i samej osobliwości. Następnie omówiono problem dostosowania SI, odnosząc się do potencjalnych skutków superinteligentnej SI dla ludzkiej cywilizacji. Wykorzystując teorię strukturacji Giddensa, podkreślono także znaczącą rolę „autorytatywnych zasobów" opartych na sztucznej inteligencji w określaniu alokacji zasobów i zapewnianiu sprawiedliwości dystrybucyjnej w społeczeństwie technoutopijnym. W artykule przeanalizowano również ideę utopii i „końca historii", a we wnioskach zasugerowano, że osiągnięcie technologicznej utopii z superinteligentną sztuczną inteligencją stanowi problem projektowania mechanizmów.

Introduction

The world economy has rapidly transformed over the past century due to the advancement of technology, combined with the evolution of institutional systems across most parts of the world. The exponential increase in the power of computing, as described by Moore's Law –where the number of transistors on a microprocessor chip, and hence its computational power, doubles roughly every two years – has driven the meteoric rise of information technology since the 1960s and the advent of the internet era in the 1990s. Recently, this exponential growth has started to falter [**Waldrop, 2016**]. Still, this increase in computational power has paved the way for the age of artificial intelligence (AI), which directly or indirectly impacts our daily lives.

A review study of AI timelines by **Roser** [**2023**] suggests that most AI researchers take the prospect of powerful AI technology seriously and do not dismiss it as a mere fantasy, with a majority of them agreeing that human-level AI or artificial general intelligence (AGI) would be developed within the next few decades. "The public discourse and the decision-making at major institutions have not caught up with these prospects. In discussions on the future of our world – from the future of our climate, to the future of our economies, to the future of our political institutions – the prospect of transformative AI is rarely central to the conversation. Often it is not mentioned at all, not even in a footnote" [**Roser, 2023**]. Even if some polls suggest that human-level AI is likely to be developed in the coming decades, in another poll of AI researchers, 92.5% of the respondents agreed that developing superintelligence is beyond the foreseeable horizon [**Etzioni, 2016**].

Figure 1. Evolution of AI Superintelligence



Source: Authors' own interpretation based on **Bostrom** [**2014**], **Tegmark** [**2017**], **Growiec** [**2022**].

The scenario where an AGI agent evolves into AI superintelligence exceeding humans' cognitive performance by a wide margin [**Bostrom, 2014**] would catapult the world to a point of no return, called technological singularity [**Ulam, 1958**]. **Growiec** [**2022**] mentions that while the economic mainstream tends to approach the notion of technological singularity with scepticism, contemporary macroeconomists such as **Gordon** [**2016**] grapple with the looming spectre of stagnation. Singularitarian thinkers, epitomised by figures such as **Kurzweil** [**2005**], focus on the expansive frontier of technological capabilities, adopting a visionary approach that spans extended timeframes. In stark contrast, pragmatically inclined stagnationist economists, such as Gordon, scrutinise realised GDP per worker over more immediate time horizons. **Growiec** [**2022**] further argues that we are living in a world of accelerating, not decelerating, change as suggested by the stagnation hypothesis.

This paper attempts to discuss the impact of AI-driven technological singularity and its possible economic repercussions. The various scenarios and concepts discussed in this paper are summed up in Figure 1, which shows the evolution from narrow AI to superintelligent AI.

The key ideas discussed in this paper are structured as follows:

1. **The Dawn of AI Superintelligence and Technological Singularity:**

First, the ideas of Artificial General Intelligence (AGI), AI Superintelligence and Technological Singularity are discussed.

2. **The AI Alignment Problem and its Consequences**:

The possible outcomes of the impact of superintelligent AI on human civilisation are discussed before understanding its economic repercussions because various dystopian outcomes are probable as well, which would make any such discussions fruitless. Potential utopian outcomes are zeroed in on where the alignment problem is already solved and the problem of control over AI as an authoritative resource takes centre stage.

3. **Envisioning Economy at Technological Singularity**

This is the concluding section of the paper and it discusses the following aspects:

3.1. Effects of superintelligent AI on the economy.
3.2. The concentration of AI power and the problem of authoritative resources.
3.3. The end of history or a new beginning?
3.4. Possibility of a techno-utopian society.

The Dawn of AI Superintelligence and Technological Singularity

The term "technological singularity" was coined and popularised by computer scientist and science-fiction author Vernor Vinge in 1983 [**Vinge, 1983**]. In 1993, he further elaborated his ideas about the posthuman era of technological singularity, when "the development of computers that are 'awake' and superhumanly intelligent" would become a reality. He wrote that "this change will be a throwing-away of all the human rules, perhaps in the blink of an eye – an exponential runaway beyond any hope of control. Developments that were thought might only happen in a million years (if ever) will likely happen in the next century. It's fair to call this event a singularity (the Singularity for the purposes of this piece). It is a point where our old models must be discarded and a new reality rules, a point that will loom vaster and vaster over human affairs until the notion becomes a commonplace" [**Vinge, 1993**].

The idea of technological singularity has only emerged within the past century and can be traced back to a few prominent thinkers. After the passing of John von Neumann in 1957, Stanisław Ulam expressed in his writings that he had engaged in a discussion with von Neumann about the rapidly advancing pace of technology and the consequential transformations in human existence. He mentioned that "one conversation [with John von Neumann] centered on the ever accelerating progress of technology and changes in the mode of human life, which gives the appearance of approaching some essential singularity in the history of the race beyond which human affairs, as we know them, could not continue" [**Ulam, 1958**].

In 1965, I. J. Good originated the concept of "intelligence explosion," rather than calling it "technological singularity." He expressed the idea of an "ultraintelligent machine" that possesses intellectual abilities far superior to any human being. Good suggested that such a machine could not only outperform humans in various intellectual activities but also have the capability to design even better machines. This would trigger an "intelligence explosion" where the machine's intelligence would rapidly surpass that of humans. Good further speculated that once the first ultraintelligent machine is created, it would be the last invention humans would ever need to make, as its superior intelligence would render further human inventions unnecessary [**Good, 1966**].

In recent times, Ray Kurzweil has popularised the idea of singularity as a point in time in the future when humans and machines will merge and create a new form of superior non-biological intelligence [**Kurzweil, 2005**]. He distinguishes "narrow AI" as systems that carry out specific "intelligent" behaviours in specific contexts. "Artificial general intelligence" (AGI) has emerged as an antonym to "narrow AI" to refer to systems that have the broad capability to self-adapt to changes in their goals or circumstances [**Goertzel, 2014**], just like humans.

Thus, AGI is a form of AI with human-like general intelligence, while AI superintelligence is a theoretical future state of AI that transcends human intelligence by a significant margin. AGI is a step on the path toward AI superintelligence, and the development of superintelligent AI is a topic of concern and debate due to the potential profound impacts it could have on society.

Recent developments such as the ChatGPT AI chatbot, which can engage in real conversations with humans and generate articles, stories, poems, and even computer codes, show remarkable progress towards attaining AGI capabilities [**Greengard, 2022**]. *Harvard Business Review* has described ChatGPT as a "tipping point for AI" [**Mollick, 2022**]. AI is pushing the boundaries of creativity by excelling in pursuits that were traditionally considered the exclusive domain of human ingenuity. Thus, the vision of technological singularity is no longer science fiction, but is becoming real with each passing year.

The AI Alignment Problem and its Consequences

The proliferation and impact of artificial intelligence and other advanced technologies can have significant civilisational consequences in the near future. These consequences can encompass both positive and negative outcomes. The key to our future lies in ensuring that a AI evolves, we are able to ensure that it aligns with human values and objectives. This is known as the "Alignment Problem of AI."

One of the earliest descriptions of the alignment problem was stated by Norbert Wiener as follows: "If we use, to achieve our purposes, a mechanical agency with whose operation we cannot efficiently interfere once we have started it…then we had better be quite sure that the purpose put into the machine is the purpose which we really desire and not merely a colorful imitation of it." [**Wiener, 1960**]. And interestingly, as we build better AI systems that are able to do hard tasks which are even beyond human capabilities, they will get integrated into all kinds of valuable applications due to economic pressures. However, if they are misaligned, they will not accomplish the tasks as we desire, resulting in unforeseen consequences [**Leike, 2022**]. The control problem, on the other hand, focuses on the challenge of controlling advanced AI systems to prevent them from taking actions that are harmful, undesirable or potentially catastrophic [**Russell, 2020**].

While there is some overlap between the two problems, they are not the same. The alignment problem is more concerned with ensuring that AI's values align with human values, while the control problem focuses on the practical mechanisms for maintaining control and avoiding undesirable consequences. Both issues are critical in the development and deployment of advanced AI systems as they relate to the safe and ethical use of AI.

It is very much a possibility that superintelligent AI will misalign from its specified human-centric objectives, leading to a variety of dystopian outcomes. The "Orthogonality Thesis" proposed by Nick Bostrom makes a convincing case that superintelligent AI can have very different non-anthropomorphic goals. He states that "any level of intelligence could in principle be combined with more or less any final goal." Thus, superintelligent AI need not care about human-like motivations based on a sense of morality, and could very well make seemingly absurd or even evil decisions from a human perspective.

A number of outcomes can occur, and a broad range of these have also been summarised (Table 1) by **Tegmark [2017]**. Most of these are dystopian in nature where humans either cease to exist or are not happy in it. The ones in which humans do exist and are happy require that the problem of control is taken care of, such that superintelligent AI does not enslave humans or wipe them out completely [**Ord, 2020**].

There appears to be a trade-off between the safety of humans and control over superintelligent AI in the utopian scenarios as discussed by **Tegmark [2017]**, where humans and superintelligence both co-exist, humans are in control (at least partially) and their safety is ensured (at least potentially):

1. Benevolent Dictator (BD) (Control: No, Safety: Yes)
2. Protector God (Control: Partially, Safety: Potentially)
3. Gatekeeper (Control: Partially, Safety: Potentially)
4. Enslaved God (EG) (Control: Yes, Safety: Potentially)

As we move along the spectrum from the Benevolent Dictator to the Enslaved God type of superintelligent AI, the autonomy of humans over its own fate increases whereas the interference of superintelligent AI decreases. Further, scenarios where AI does not fully conform to human objectives and acts in a manner to mask its own existence and prevents another superintelligence to be developed are inconsequential to our discussion, as the knowledge of even their existence will be hidden from humans.

Humans might become existentially safe under a benevolent and dictatorial superintelligent AI taking care of the needs of the entire civilisation, but this can lead to a loss of existential meaning and suppression of individual liberty. Such a scenario is mentioned as the "purpose problem" by Yuval Noah Harari in his book *Homo Deus* [**Harari, 2017**]. It arises from the fact that humans are hardwired to seek meaning and significance in their lives. Without a clear sense of purpose, individuals may feel lost, disoriented or even disillusioned. This is not something new in human history, when various forms of oppressive regimes such as feudalism, slavery, the caste system, apartheid, and colonialism have had the same effect. Harari suggests that in the future, as our traditional belief systems erode, we may need to find new ways to navigate this "purpose problem" and create meaningful lives for ourselves. This might also require designing purpose-driven incentive structures for humans, which superintelligent AI might be able to accomplish. In a more extreme scenario, superintelligent AI could also act as a zookeeper of human specimens, where all sense of freedom and purpose is lost.

The Enslaved God scenario, on the other hand, though allowing for increased human control, can also lead to self-destruction of the human civilisation due to competing interests in the worst-case scenario. The important theme that therefore emerges is that *even if the alignment problem is solved and we are headed for a techno-utopian society, the control over the access and usage of superintelligent AI resource remains a significant problem afterwards*. If humans are allowed full autonomy over their choices in the way they use superintelligent AI, it is essential that adequate governance institutions and checks and balances are in place to avoid unintended or harmful consequences arising out of human discretion.

Envisioning Economy at Technological Singularity

Effects of superintelligent AI on the economy

A variety of technological advancements, such as AI, robotics, quantum computing, nanotechnology, gene editing, the transition from fossil fuels to carbon-neutral renewable resources, blockchain, and 3D-printing, will combine to disrupt the way economists think about three fundamental problems in an economy: what to produce, how to produce, and for whom to produce.

The issue of the impact of automation on productivity and labour demand is an old one, and has already been discussed by many prominent economists. The most straightforward impact of the proliferation of AI and robotics is that it can significantly enhance productivity and alter the labour market equilibrium. Broadly, it can either augment human labour ("productivity effect") or completely replace it ("displacement effect"), as noted by **Acemoglu and Restrepo [2018]**. The displacement effect can reduce the demand for labour, wages and

employment, as the increased productivity outpaces the overall labour demand. The productivity effect is a countervailing force that pushes against the displacement effect and could therefore increase the labour demand.

In 1930, John Maynard Keynes published a short essay entitled "Economic Possibilities for Our Grandchildren," in which he prophetically claimed that "the economic problem may be solved, or be at least within the sight of solution, within a hundred years" [**Keynes, 2010**: 326]. He also discussed the problem of unemployment due to technological advancement, and popularised the term "technological unemployment," which, according to him, would only be a "temporary phase of maladjustment" in the pursuit of "mankind solving its economic problem" [**Keynes, 2010**: 325].

Keynes rightly pointed out that the most pressing problem for the human race and the entire biological kingdom, since the beginnings of life on earth, has been the problem of subsistence. Keynes further theorises that the pace at which we can reach the point of "economic bliss" will be governed by four things: control of overpopulation; avoidance of wars and civil unrest; development and use of scientific temperament and means in society wherever required; and the rate of capital accumulation, which according to him "will easily look after itself, given the first three" [**Keynes, 2010**: 331].

Keynes was right about the rapid pace of economic growth and the rise in the standard of living for most of the world, but his predictions about work and leisure are far off from the economic reality prevailing in our times. He thought that people, when liberated from "absolute needs" [**Keynes, 2010**: 326], would work no more than 15 hours a week and spend the rest of the time on leisure and culture, i.e., non-economic pursuits. Also, he did not delve into the problems of distribution and inequality. He only cursorily mentioned that "avarice and usury and precaution must be our gods for a little longer still. For only they can lead us out of the tunnel of economic necessity into daylight" [**Keynes, 2010**: 331]. Thus, he referred to the ills of capitalism as something that must be endured until human civilisation reaches the point of "economic bliss" when all the problems of subsistence have been solved, but he did not explain how capitalism would evolve into such an economic utopia.

As AI evolves to its ultimate superintelligent form in the future, humans should be able to evolve the structure of capitalist economies as well with the ever increasing power of AI. From a theoretical viewpoint, a future can be conceived where all human labour has been replaced by artificially intelligent agents and the entire chain of the economy is managed by superintelligent AI, which efficiently produces and allocates resources according to new incentive structures of society. The dawn of technological singularity could allow humans or superintelligent AI to become the ultimate beekeeper of intelligent machines, which would manage the production of goods and services.

Given these possibilities, the first two fundamental problems of economics, i.e., what to produce and how to produce, can be optimised and taken care of. That's assuming that superintelligent AI can augment production processes based on the principles of balancing efficiency and equity for society as a whole, thus allowing for the existence of utopian scenarios as discussed earlier. Then the central problem which will demand the attention of all humanity will be distribution, i.e., for whom to produce. One problem is the distribution of material resources, but more importantly, it is the distribution and governance of AI resources that decides if an equitable distribution of the former can be achieved.

Control over superintelligent AI as an authoritative resource

Anthony Giddens [**Giddens, 1995**] has distinguished "two major types of resources that enter into structures of domination: those that are involved in the dominion of human beings over the material world (allocative resources) and those involved in the dominion over the social world itself (authoritative resources)." Authoritative resources allow agents to control persons, whereas allocative resources allow agents to control material objects. The effects of technological singularity would fundamentally alter the relationship of humans with allocative and authoritative resources.

The control of superintelligent AI could act as the ultimate authoritative resource, reshaping power structures within human society and influencing the masses to support such control. This takes the Marxist idea of ownership of the means of production to its extreme. While unevolved AI resources and simple machines may be seen as means of production in a traditional sense, ensuring human control of superintelligent AI is equivalent to the "Enslaved God" scenario. This kind of control could make the masses unaware of the existence of superintelligent AI, potentially leading to an Orwellian future.

Even now, it can be seen that some forms of advanced AI are accessible only to those who own or can afford them. Therefore, it is plausible that superintelligent AI could eventually be controlled by a select few individuals. It is highly unlikely that such an outcome would lead to an equitable distribution of allocative resources. The development of superintelligent AI must ensure that its control within society aligns with human welfare, while also ensuring distributive justice, for example the Rawlsian principles of liberty and equality within society [**Rawls, 1971**]. The process must also ensure that all individuals are fully empowered to live the lives they value [**Sen, 1980**], in line with the capability approach developed by Amartya Sen, Martha Nussbaum, and others [**Robeyns, 2005**].

If we consider the Benevolent Dictator and Enslaved God scenarios as the only cases of interest, then it is highly likely that the EG scenario would result in an oppressive regime, leading to the exploitation of the masses by the few who control the superintelligent AI resource. Thus, the developmental phase of AI, where it transitions from AGI to superintelligent AI, is perhaps the most critical phase determining welfare in a post-singularity society. Some of the mechanisms have been discussed by **Bostrom** [**2014**] to solve the control problem classified as capability control methods and motivation selection methods. A combination of these methods could be used to ensure that superintelligent AI does not fall into the control of a few people. Further, formal verification proofs could be used to ensure that AI systems behave as intended, at least when looking at the broad objective of avoiding concentration of control within a small group of individuals.

The end of history or a new beginning?

The idea of the "end of history" can be traced back to the German philosopher Georg Wilhelm Friedrich Hegel, who proposed in his seminal work "Phenomenology of Spirit" that human history was progressing towards a state of absolute freedom and rationality [**Hegel, 2018**]. He believed that human societies were evolving towards a state of self-realisation, and that the ultimate goal of history was the realisation of a free and rational society. Hegel saw the development of the modern state as the ultimate manifestation of this process and believed that it represented the end of history.

Francis Fukuyama, drawing on Hegelian ideas, used the term "end of history" in a more specific and controversial sense in his 1989 essay "The End of History?" He argued there that the fall of the Soviet Union and the triumph of liberal democracy had signalled the end of history. He suggested that liberal democracy had become the final form of government, and that there would be no further significant changes in the political system. Fukuyama saw the end of the Cold War as the culmination of a long process of historical development and argued that liberal democracy had become the only viable political system [**Fukuyama, 1989**].

However, Fukuyama's thesis has been criticised by many philosophers, who argue that history is an open-ended process, and that there is no endpoint to human development. Some critics have suggested that Fukuyama's thesis is based on a simplistic view of history that fails to consider the complexities of human societies. They argue that Fukuyama's thesis ignores the role of culture, religion, and social movements in shaping human societies, and that it is based on an overly deterministic view of history [**Sandel, 1996**].

The concept of the "end of history" has also been explored in the history of economic thought. Karl Marx, the German philosopher and economist, argued that history was a dialectical process, and that capitalism was a necessary stage in the historical development of human societies. Marx believed that capitalism would eventually lead to its own downfall, and that a new form of society, communism, would emerge as the final

stage of historical development. He saw communism as the end of history, where there would be no more class struggle or social conflict [**Engels, 2019**].

However, Marx's thesis has been criticised by many economists, who argue that capitalism has shown remarkable resilience, and that it is unlikely to be replaced by communism. They suggest that the market economy has proven to be a highly efficient system of resource allocation, and that it is unlikely to be surpassed by any other economic system. They also argue that communism has been discredited by its failure in the Soviet Union and other socialist countries [**Friedman, 1962**].

If AI were to achieve superintelligence – surpassing human cognitive capabilities in every aspect – it might catalyse a revolutionary shift in the socio-political landscape. Potential consequences could include the establishment of an entirely new paradigm where traditional human-driven history takes a backseat to the autonomous decision-making and problem-solving abilities of the superintelligent entity.

The prospect of AI superintelligence triggering the end of history is also intricately linked to the Alignment Problem and the Problem of Control, as can be seen from the various scenarios in Figure 1. A well-aligned AI might contribute to solving complex global challenges, accelerating scientific progress, and enhancing societal well-being, thus heralding a new beginning for the human civilisation. However, if the alignment is not carefully addressed, the consequences could range from unintended negative impacts to a radical reshaping of human history.

Possibility of a techno-utopian society

Keynes' concept of "economic bliss" as discussed earlier is similar to the ideas of Utopia as envisioned by various thinkers in the past. The word "utopia" is derived from a Greek term meaning "no place." It was first coined by Sir Thomas More in 1516 and usually refers to an ideal society which is usually stationary in its economic aspect [**Claeys, 1989**].

The idea that technology could usher in an era of utopia is not new and is referred to as technological utopianism [**Segal, 2005**], where machines alone are the means to achieve an ideal utopian society. Thus, the role of technological singularity in realising an economic utopia is in line with the techno-utopian ideology. It could be that the end of history as "the end-point of mankind's ideological evolution" [**Fukuyama, 1989**] would coincide with the epoch of technological singularity, giving rise to a techno-utopian society. But Fukuyama's idea that Western liberal democracy would act as the final form of human government [**Fukuyama, 2006**] and be the final answer to the problem of distribution of "authoritative resources," as mentioned by Giddens [**Giddens, 1995**], does not take into account the governance structures and institutions that superintelligent AI could create and enforce in the future.

The political structure of liberal democracy, coupled with capitalism as the predominant economic model championed by the West, emerged triumphant at the end of the 20th century with the fall of the USSR. But many countries have thrived economically without being a liberal democracy. China stands as a prime example, with its unique blend of authoritarian governance and market-oriented economic policies driving rapid growth and development. Many Middle Eastern countries, with their autocratic rule, have transformed themselves into a modern and diversified economy, fuelled by oil wealth. Similarly, a future governed by or with the help of superintelligent AI can lead to divergent political outcomes, even if economic well-being remains the predominant criterion as part of its alignment. Thus, creating a technological utopia entails more than just developing advanced technology; it also presents a mechanism design problem that requires careful consideration of how technology will be deployed and how societal structures will be transformed. Such a future juncture would necessitate a reorientation of human endeavours and socio-political institutions, reframing the distribution of "authoritative resources" as a mechanism design challenge.

According to **Bostrom** [**2014**], the design of a technological utopia must be based on the principle of fairness. He argues that the creation of superintelligent AI could lead to the concentration of power and wealth in the hands of a small group of individuals, potentially leading to catastrophic outcomes. Therefore,

he suggests that we must design AI systems that are aligned with human values and goals, rather than allowing them to pursue their own objectives. Another challenge is designing systems that promote sustainability and avoid an ecological catastrophe. Technological progress has the potential to exacerbate environmental problems, such as climate change and resource depletion, if not deployed carefully. Therefore, it is essential to develop technologies that are environmentally sustainable and to create economic and social systems that prioritise sustainability over short-term profits.

As **Sen [1999]** argues in "Development as Freedom," the goal of economic development should be to expand the capabilities and freedoms of all individuals, rather than simply increasing GDP. This principle is especially relevant in the context of a technological utopia, where the goal should be to use technology to improve the well-being and quality of life of all individuals, rather than just the privileged few.

Finally, designing a technological utopia also requires careful consideration of the values and ethics that should guide our technological development. As technology becomes increasingly powerful, it has the potential to shape society in profound ways, potentially leading to the creation of new forms of inequality and oppression. Therefore, it is essential to develop a framework for thinking about the ethical implications of technology and to ensure that technological development is guided by values such as social justice, equality, and human dignity [**Harari, 2017**].

Economists will need to think beyond conventional economic paradigms and devise new methods for structuring incentives and distribution mechanisms for human civilisation. Assuming humans remain distinct from, and not absorbed into, superintelligent AI in a transhumanist scenario, and are not rendered extinct, the control of AI itself will determine the distribution of wealth among human labour. If AI is treated as capital, the rise of highly efficient machine intelligence replacing human labour could drastically reduce wages. With the factor share of capital nearing 100% due to a surge in global GDP resulting from an intelligence explosion and technological advancements, total income from capital would skyrocket [**Bostrom, 2014**]. Therefore, the transition towards a superintelligent AI-based economy must incorporate some form of social security measures during the transition period.

Factor payments in a labour-free world may take on new forms. For example, there could be payments for the use of AI algorithms, data, and intellectual property rights, which would become valuable assets in an automated economy. Prices in a market economy are often determined by the costs of production, including labour and capital costs. Without factor income, there would be no direct cost to consider when setting prices. Instead, prices might be influenced by factors such as resource scarcity, technology, and demand. Governments may need to introduce policies such as universal basic income (UBI) or other social safety nets to address the loss of labour income and prevent widespread economic hardship. These policies would involve redistributive payments to ensure that people can meet their basic needs despite the lack of traditional employment opportunities. Economists and policymakers would need to develop new theories and frameworks to understand and manage the dynamics of a labour-free economy. The question of whether and how wealth could be rationalised in a post-singularity world is open to debate.

By designing AI systems that promote fairness, sustainability, and ethical values, we may be able to move closer to the ideal of a perfect world. However, achieving this goal will require significant investment in research and development, as well as a willingness to challenge existing power structures and systems of inequality. While the creation of a technological utopia is a challenging and complex problem, the potential benefits of such a society make it a goal worth pursuing.

References

Acemoglu D., Restrepo P. [2018], The Race between Man and Machine: Implications of Technology for Growth, Factor Shares, and Employment, *American Economic Review*, 108 (6): 1488–1542.

Bostrom N. [2014], *Superintelligence: Paths, dangers, strategies*: xvi, 328, Oxford University Press, Oxford.

Claeys G. [1989], Utopias, in: Eatwell J., Milgate M., Newman P. (eds.), *The Invisible Hand*: 270–277, Palgrave Macmillan, London, https://doi.org/10.1007/978-1-349-20313-0_36.

Engels F., Marx K. [2019], The Communist Manifesto, in: *Ideals and Ideologies*, 11th ed., Routledge, New York.

Etzioni O. [2016], *No, the Experts Don't Think Superintelligent AI is a Threat to Humanity*, MIT Technology Review, Retrieved December 12, 2023, from https://www.technologyreview.com/2016/09/20/70131/no-the-experts-dont-think-superintelligent-ai-is-a-threat-to-humanity.

Friedman M. [1962], *Capitalism and Freedom*, University of Chicago Press, Chicago.

Fukuyama F. [1989], The End of History?, *The National Interest*, 16: 3–18.

Fukuyama F. [2006], *The End of History and the Last Man*, Simon and Schuster, New York.

Giddens A. [1995], *A Contemporary Critique of Historical Materialism*, Stanford University Press, Stanford.

Goertzel B. [2014], Artificial General Intelligence: Concept, State of the Art, and Future Prospects, *Journal of Artificial General Intelligence*, 5 (1): 1–48, https://doi.org/10.2478/jagi-2014-0001.

Good I. J. [1966], Speculations Concerning the First Ultraintelligent Machine, in: Alt F. L., Rubinoff M. (eds.), *Advances in Computers*: 31–88, Elsevier, https://doi.org/10.1016/S0065-2458(08)60418-0.

Gordon R. J. [2016], *The Rise and Fall of American Growth: The U. S. Standard of Living Since the Civil War*, Princeton University Press, Princeton, https://econpapers.repec.org/bookchap/puppbooks/10544.htm.

Greengard S. [2022], *ChatGPT: Understanding the ChatGPT AI Chatbot*, eWEEK, Retrieved January 31, 2023, from https://www.eweek.com/big-data-and-analytics/chatgpt.

Growiec J. [2022], *Accelerating Economic Growth: Lessons From 200,000 Years of Technological Progress and Human Development*, Springer International Publishing, https://doi.org/10.1007/978-3-031-07195-9.

Harari Y. N. [2017], *Homo Deus: A Brief History of Tomorrow*, HarperCollins, New York.

Hegel G. W. F. [2018], *Hegel: The Phenomenology of Spirit*, Oxford University Press, Oxford.

Keynes J. M. [2010], Economic Possibilities for Our Grandchildren, in: *Essays in Persuasion*: 321–332, Palgrave Macmillan, Basingstoke.

Kurzweil R. [2005], *The Singularity is Near: When Humans Transcend Biology*, Viking Press, New York.

Leike J. [2022], What is the alignment problem? [Substack newsletter], *Musings on the Alignment Problem*, https://aligned.substack.com/p/what-is-alignment (accessed on 7.07.2023).

Mollick E. [2022], ChatGPT Is a Tipping Point for AI, *Harvard Business Review*, https://hbr.org/2022/12/chatgpt-is-a-tipping-point-for-ai (accessed on 31.01.2023).

Ord T. [2020], *The Precipice: Existential Risk and the Future of Humanity*, Hachette Books, New York.

Rawls J. [1971], *A Theory of Justice: Original Edition*, Harvard University Press, Cambridge, https://doi.org/10.2307/j.ctvjf9z6v.

Robeyns I. [2005], The Capability Approach: A Theoretical Survey, *Journal of Human Development*, 6 (1): 93–117, https://doi.org/10.1080/146498805200034266.

Roser M. [2023], *AI timelines: What do experts in artificial intelligence expect for the future?*, Our World in Data, https://ourworldindata.org/ai-timelines (accessed on 31.01.2023).

Russell S. [2020], *Human Compatible: Artificial Intelligence and the Problem of Control*, Penguin Publishing Group, New York.

Sandel M. J. [1996], *Democracy's Discontent: America in Search of a Public Philosophy*, Harvard University Press, Cambridge.

Segal H. P. [2005], *Technological Utopianism in American Culture*, Syracuse University Press, Syracuse, https://digitalcommons.library.umaine.edu/fac_monographs/138.

Sen A. [1980], Equality of What?, *The Tanner Lecture on Human Values*, I: 197–220.

Tegmark M. [2018], *Life 3.0: Being Human in the Age of Artificial Intelligence*, Penguin Books, London.

Ulam S. [1958], John von Neumann 1903–1957, *Bulletin of the American Mathematical Society*, 64 (3): 1–49, https://doi.org/10.1090/S0002-9904-1958-10189-5.

Vinge V. [1983], First Word, *Omni*, 5 (4), https://www.isfdb.org/cgi-bin/title.cgi?120008 (accessed on 6.03.2023).

Vinge V. [1993], The Coming Technological Singularity, *Whole Earth Review*, Winter Issue.

Waldrop M. M. [2016], The chips are down for Moore's law, *Nature News*, 530 (7589): 144, https://doi.org/10.1038/530144a.

Wiener N. [1960], Some Moral and Technical Consequences of Automation, *Science*, 131 (3410): 1355–1358, https://doi.org/10.1126/science.131.3410.1355.

Appendix

Table 1. Properties of AI Aftermath Scenarios

Scenario	Superintelligence exists?	Humans exist?	Humans in control?	Humans safe?	Humans happy?	Consciousness exists?
Libertarian Utopia	Yes	Yes	No	No	Mixed	Yes
Benevolent Dictator	Yes	Yes	No	Yes	Mixed	Yes
Egalitarian Utopia	No	Yes	Yes?	Yes	Yes?	Yes
Gatekeeper	Yes	Yes	Partially	Potentially	Mixed	Yes
Protector God	Yes	Yes	Partially	Potentially	Mixed	Yes
Enslaved God	Yes	Yes	Yes	Potentially	Mixed	Yes
Conquerors	Yes	No	–	–	–	?
Descendants	Yes	No	–	–	–	?
Zookeeper	Yes	Yes	No	Yes	No	Yes
1984	No	Yes	Yes	Potentially	Mixed	Yes
Reversion	No	Yes	Yes	No	Mixed	Yes
Self-Destruction	No	No	–	–	–	No

Source: Tegmark [2018: 160–161].