Philosophy and Consciousness in the Future: Cyborgs and Artificial Intelligence waiting for Immortality

Abstract

Natural sciences and technologies place artificial intelligence, robotics and cyborgs at the centre of human attention. However, virtual and augmented reality and the unthinkable possibilities of the future media and communication between individuals and social groups might be deeper and broader than we think, and evolve in forms we have not hoped for. If philosophy wants to focus on the essence of scientific-technical age, it must reflect on the very foundations of computer-operational thinking, because artificial intelligence has already started to develop its own consciousness and may, in the future, devise a strategy of development beyond man and without man.

Key words: future, philosophy, cyborgs, communication, media, consciousness, artificial intelligence.
The missing link between animals and the real human being is most likely ourselves

Konrad Lorenz

Questions such as “when will robots acquire consciousness?” or “can machines replace the human brain?” or “at what point will artificial intelligence overpower human and what will happen then?” were once reserved mainly for science fiction and as such often outside the interest of philosophy. SF authors such as Arthur Clarke, Isaac Asimov, Stanislaw Lem, or Ray Bradbury were perceived as the heralds or pioneers in discussions about the future or even its “contours”. Modern discoveries and insights, especially since the end of the past century, have confirmed some of their visionary solutions, which is why questions about the future are now increasingly considered in relation to the present quantity and quality of scientific-technological development. It has become clear that the future is already here, among us, and that it must be seen in its essence if we are to direct it towards human ends.

The future has been gradually accepted as an area of new research in which the human mind will re-examine not only its rational, but also its ethical abilities. Michio Kaku emphasizes that predicting the future has long been the subject of human desires and dreams, from ancient history to the latest technologies. Regardless of these dreams from the ancient times to the SF authors, Kaku has seen it as one of the essential features of human nature and human civilization. For him, its key feature is to “translate thinking into reality” and to fight the prejudice that it is impossible to predict more closely the world, let’s say, in a hundred years. Despite the obstacles, he considers thinking about the future as an indispensable part of scientific work, and this kind of philosophy simply needs to be fostered and developed. Therefore, it is also important to “closely examine the work of the leading scientists who are building prototypes of the technologies that will change our future... The prototypes of all these technologies already exist.”

When thinking about the future within the development of science and technology, it is important to establish an “inevitable condition,” namely a higher level of awareness among people, scientists and/or visionaries about the way in which the mind can reach assumptions about what is to come. Kaku seems to be pessimistically concluding that predictions about the turn of the 22nd century are nevertheless a discouraging task, although not because of some limitation of human thought or the methods of philosophizing-about-the-future.

13 Kaku refers to persons such as Jules Verne, who in his literary works – precursors of the later SF novels – could “see” with great precision all the details of the future development of technology and what it would cause. Similar were the projects of Leonardo Da Vinci, and one of the common features of the two visionaries was their trust in science and an extrapolation of the achievements of their own era into what they believed – and predicted – would almost certainly follow. It is interesting to see how one of today’s leading scholars emphasizes the importance of vision in both of them, the writer and the painter, noting that, for him, they – like many other visionaries – are not just representatives of their “genres”, but rare individuals who seriously, responsibly, and consciously thought about the future.

The problem is, first and foremost, in the rapid and increasing pace of change, the sudden leaps in the development of science and technology that are difficult to follow, especially when it comes to the humanist understanding of the consequences of this “ever-accelerating science,” as well as the uneven development of the society and human psychology. In the last few decades alone, practically more scientific knowledge has been collected than in all of the past, and such dynamics of scientific breakthroughs in technology and the human understanding of these changes will surely accelerate even more by the end of this century.

Any even remotely philosophical understanding of this accelerated change, of its rhythm and dynamics, conceals a sort of absurdity. The classic way of understanding and thinking, in which analysis and synthesis (inside and outside) yield a specific “category”, a closed system that interprets something and provides a solid basis for further thinking, has lost some of its characteristics. Future thinking, as well as thinking (about) the future, has become increasingly elusive when it comes to a compact definition and constitution of long-term determinants to be used in classical observation in terms of “subject and object”. I am specifically referring to the accelerating change and especially the breadth of discoveries and insights regarding cybernetics, virtual reality, cyborgization in the application of the most advanced technologies, and finally what we call “artificial intelligence,” which we even fear in public debates.15

A similar concern is partly voiced in public, even among the professionals and scientists, in relation to the new issues of postmodern reality and environment, such as cyberspace, cyberbody,16 and anything else that focuses on the possibilities of different forms of human existence, i.e. the coexistence and permeation of organic and inorganic life, all the way down to the quantum and bio-computers, or implanting various “gadgets” in one’s body, from limbs and internal organs to the brain itself.

Mike Featherstone and Roger Burrows have pointed out that “those people who lived a few decades before the Renaissance were unable to see it coming. This belief that something completely new could be just around the corner, that humankind still has an open future, is one which has been challenged by postmodern theory with its attacks on the modernist metanarratives of progress and “the new”. They have indicated the ambivalent feelings, prognoses, and conclusions related to the likewise increasing postmodern pessimism based on their analysis of everyday events around global

---

15 It would take more space to explain how the present and future of the so-called artificial intelligence has outgrown the original analyses dealing with “machines”, “robotics”, and androids. In perhaps less than a decade, instead of the aforementioned “frame” of the topics, scientists (as well as general public) have focused on algorithms, cyborgs, humanoid robots that we will hardly be able to distinguish from human beings, conversations with programs that are already about to pass the Turing Test, and artificial intelligence that is able to produce further AI based on its own experience. Not to mention the fact that the gradual “disappearance” of the digital age (to which part of the older population has not yet even become accustomed) is being proclaimed in favour of an “age of artificial intelligence.”

16 For a broader discussion, including trans-humanist premises and reflections “beyond” the usual ones in the humanities, see also: Nenad Vertovšek and Ivana Greguric-Knežević, “Filozofija budućih kiberprostora i transhumanistička stvarnost” [The philosophy of future cyberspaces and the trans-humanist reality], Filozofska istraživanja 149 (1/2018), 99-116.

wars, climate change, inequality, and conservative brakes hindering the advancement of science and technology, while at the same time there seem to surge new utopian understandings about a turning point leading to the future and solving almost all accumulated problems, even those that are yet to emerge.

A good mental exercise for questioning human prediction skills are musings from the 1960s, which did not envision computers in daily life 30 years later, preferring prognoses about robotic development?! Some, like Mark Poster, believed that at the turn of the millennium we would be in some sort of a historic situation, same as at the time of the emergence of urban trading culture in feudalism, since the end of the twentieth century brought about unpredictable upheavals in media development, from the advent of the Internet to the emergence of social networks.

Featherstone and Burrows have added something that is crucial to our consideration of consciousness and artificial intelligence: “It is not just the possible reconstitutions of social life and culture which interest us in this journal, it is the impact of these changes on the body, too. It is here that developments of technology point towards the possibilities of post-bodied and post-human forms of existence. If the development of technology has entailed a process of the extension of the body and bodily functions to enable us to control the environment more efficiently, it offers the ultimate possibility of the displacement of the material body from the confines of its immediate lived space.”

Questions about the new forms of human existence thus run parallel to those on “new consciousness”, which would not only relate to our present human perception, to understanding the world around us with our existing senses and mind, but also to consciousness that would include not only the organic life we know, but also various non-human (which is different than inhuman) elements of existence and being.

A deeper reflection on post-corporeality and assessment of the post-human aspects that would be at the heart of a future philosophy, and not only of the history of man and the essence of the human species, is actually part of the future “story” of the coexistence and co-operation of the known man and that which could arise from enhancing the general and specific human traits and potentials. Also, the human mind will probably have to come to terms with the need to master some new segments of the future science and technology, first of all in connection with changes in the concept of the corporeal and the real, and then also with what we mean now (and what we will mean in the future) when we talk about “artificial intelligence” (quotation marks are here because the present concepts and definitions are likely to be insufficient, too narrow, and too imprecise for

---

18 It is especially interesting that these two authors foresaw from their own time (the beginning of the new millennium) that in twenty years (which is our time) computers, information technology, and electronic media would surely play the central role in predictions. Nevertheless, they wondered: “Is there still the disturbing possibility that we could have missed something which will emerge and have crucial significance?” Is artificial intelligence that “something” and is the true arrival of AI in our worlds still due, are we partly late in our predictions, and when will the greatest and most powerful change happen? The questions are, of course, innumerable, one leading to another due to so many different answers.

what is coming). It becomes even more difficult when we try to imagine the “consciousness” of what we call AI today, and the question is what we will call it in the decades to come, and it sounds again like SF when we take into account the possibility that AI may have its own proper names and definitions to refer to itself!

Non-human issues are related not only to cyborgs as trans-human beings, but also to robots as the post-human mechanical descendants who are becoming increasingly significant participants in scientific and technical realities. Tendentially, they will take over the entire physical and mental work of man and replace biology with technology. Thus, “we seem determined to give human qualities to objects and content to treat each other as things.” Without artificial intelligence, a robot is just a mechanical machine that is switched on and off according to certain tasks it performs in a given program. By incorporating artificial intelligence equal to human, or super-intelligence embedded in the future, robots will be able to pursue strategic goals of intelligence or super-intelligence that by no means need to align with human goals. Since cybernetic science and technology measure the entire life of a human being against the binary code of a question of robotic consciousness and empathy with respect to the human parents, artificial intelligence will solve it by incorporating software that mimics the appropriate emotions and ethical principles. Time will show whether the overall behaviour of artificial intelligence and super-intelligence in our mechanical offspring, who will autonomously create its own counterparts, will be analogous to human and in coexistence with man.

Another direction for the application of artificial intelligence is to create a networked virtual world that allows for a disembodied journey through infinite space and time. Today, biological man is in a relationship and in line with the virtual world, which is not a copy of the human one, but a world based on its own foundation set by artificial intelligence. With the advent of human cyborgization, the network becomes a part of us – within us. When encountering a computer, a part of our mind seems to become part of it – according to Turkle, it indeed becomes that – another self, a mirror of the mind. In this encounter, where man sees himself differently, the boundaries are blurred between the real world and the virtual one, which man enters and exits at his will, reflecting himself in the machine like in a mirror.

Thus, through computers, we are witnessing “a shift in how we create and experience our own identities,” whereby “our new devices provide space for the emergence of a new state of the self, itself, split between the screen and the physical real, wired into existence through technology.”

---

20 Even now most humans and the human “public” completely imprecisely and inaccurately speak about “machines” and “robots” largely as some kind of mechanical beings, technologically adjusted and collapsible. This has also defined the AI development, as it is once again believed that humanoid robots will receive intelligent circuits. The possibilities offered by the quantum computational future and the already mentioned compounds of biological entities and neural-computational are less often considered, probably because they are difficult to imagine for most.


22 S. Turkle, Alone Together, xiv.

23 Ibid., 16. Cf. I. Greguric, Kibernetička bića u doba znanstvenog humanizma: Prolegomena za kiborgoetiku [Cybernetic...
This simulation of life, immersion of the real into the self-standing virtual, prepares our mind and biological body for a coexistence with cyber-beings, which will lead to the submerged human being without a self, or a connection with the biological body explaining the meaning of its existence with the help of artificial intelligence.

Philosophy faces the challenge of thinking about the relationship between identity and the physical body in reality and in virtual reality as the human mind will adapt to new aspects of virtuality with the changes in corporeality in a cyborgized reality, and then also to artificial intelligence or super-intelligence, which will both reveal new dimensions of biological reality and set up a new, cybernetic reality.

Philosopher David Chalmers has catalogued as many as 20,000 articles and texts on human consciousness and self-awareness, noting that there is actually no real consensus about it. As Michio Kaku says, “Consciousness, unfortunately, is a buzzword that means different things to different people. Sadly, there is no universally accepted definition of the term. I personally think that one of the problems has been the failure to clearly define consciousness and then a failure to quantify it. But if I were to venture a guess, I would theorize that consciousness consists of at least three basic components:

1. sensing and recognizing the environment
2. self-awareness
3. planning for the future by setting goals and plans, that is, simulating the future and plotting strategy.”

In this mode of measurement and the “1 to 10” self-awareness scale, Kaku has identified parts of sensing and recognizing the environment even in simple machines such as the thermostat, but this is actually all at the bottom of the scale. He notes that the current status and possibilities of artificial intelligence may be located somewhere at the cockroach level, while after the probably rapid rise in the evolution of recorded intelligence and mind, intelligent robots and cyborgs might in the future be compared to dogs, cats, then to primates, human apes and eventually – humans...

A noteworthy account is related to the development of computer systems, social networks, and interrelations between human communities in the future, with algorithmic capabilities and profiling, as well as environmental recognition not only in biological and physical terms, but also in the socio-psychological understanding of the human qualities of communication, expression, and behaviour.

---


Mark Zuckerberg, the founder of Facebook, established in mid-2017 that “the sociopolitical upheavals of our time – from rampant drug addiction to murderous totalitarian regimes – result to a large extent from the disintegration of human communities... He promised that Facebook will lead the charge to rebuild these communities and that his engineers will pick up the burden discarded by parish priests. ‘We’re going to start rolling out some tools’, he said, to ‘make it easier to build communities.’ He further explained that... ‘We started building artificial intelligence to do this. And it works. In the first six months, we helped 50 per cent more people join meaningful communities.”

This kind of intoxication and laudations addressed at the establishment of online systems and structures to actually continue to encourage the growth of some kind of online existence at the expense of the discredited offline existence is just one aspect, somewhat reshaped, of unconditional joy and confidence in the future. According to these views, artificial intelligence will not be “liberated”, but further “exploited” for (even very ordinary and frivolous) human interests, the egoistic goals of particular groups and lobbies, those who, given their technological superiority, often consider themselves “elected”.

Thus also Harari, in his otherwise supreme trilogy, while speculating about the present and future of humanity, and while invoking optimistic signs of the future, nevertheless seeks to critically embrace the “universal” human solution. One must be careful regardless of and even when these involve a kind of “selfless help” offered to humans by AI, which will be more intelligent and many times better than today. For, let’s add, the dangers and opportunities of misusing science and technology (historically experiential) lie less in artificial intelligence and more in human non-intelligence.

Another author, David Gamez, when speaking of man facing his own quest of consciousness, where we could probably also find a foothold for argument about the special place of mankind in history and nature, begins with the statement: “We cannot imagine things that cannot become conscious. We cannot imagine an invisible physical world that has none of the properties we encounter in our bubbles of experience. We can imagine large brains, small brains, blue brains, green brains, brains made of cheese, and so on. But the physical brain cannot be imagined as it is in itself, outside all bubbles of experience.” Gamez warns that it is difficult and practically impossible to reach “pure” reality, reality in itself, beyond the mental assumptions and categories that we create by ourselves, even when we try to think about our own brains. Neural activity is hard to imagine by itself, we need display and visualization systems.

---

25 Yuval Noah Harari, *21 Lessons for the 21st Century* (London: Jonathan Cape, 2018), e-book, 176. Although Harari praises such initiatives in which artificial intelligence is involved in building general well-being, Zuckerberg’s statements still need to be considered in relation to his statements that the goal of creating Facebook was to actually enter its users’ psychological lives and gather information and personal data. Affairs such as Cambridge Analytics, that is, collecting and selling data on tens of millions of users, and attempts creating so-called crypto-currencies via social networks have overshadowed the external outpours of benevolence by Facebook’s founder (whose popularity has, by the way, declined with the younger generations), and the search for the general well-being, inequality, and exploitation of not only states, but entire regions, one should say, is not achieved, after all, by building up significant communities (for whom, how, and what?) of this type.

More precisely: “We cannot imagine the invisible physical world. So thought experiments and imagination cannot be used to study the relationship between invisible physical brains and conscious experiences. They can only be used to study the relationship between our conscious experiences of brains and other conscious experiences. As brain-scanning technology improves we will find it easier to make imaginative transitions between conscious experiences of brain states and other conscious experiences.”

The key moment, it seems, in human reflection on consciousness is the question of whether anyone else but a human can really – think like a human. Although we have been exploring and finding thought activities in animals, aspects of intelligence in certain species and/or specimens of chimpanzees, dolphins, and elephants, we have been reluctant to take a step further towards artificial intelligence and “machines”. With some hesitation, we often ask questions about how and to what extent AI can progress: “To actually create a technical model of full blown, perspectively organized conscious experience seems to be the ultimate technological utopian dream. It would transpose the evolution of mind onto an entirely new level […]. It would be a historical phase transition. […] But is this at all possible? It certainly is conceivable. But can it happen, given the natural laws governing this universe and the technical resources at hand?”

Gamez, as well as authors like Kaku, Featherstone, Wiener, and others ask numerous questions concerning artificial intelligence, often even “in reverse”, especially when talking about the so-called point of emergency at which AI will transcend and surpass human intelligence. Most scientists who are researching the “point of emergence” today disagree about the time when this could happen and predictions range from 20 years to over a century from now. But everyone agrees that it will most certainly happen.

A new or additional hypothesis to this debate is certainly its “twist”, as we try not to look at the “issue” solely from the “human perspective, especially when we want to answer the question of whether machines can think, and how to view intelligence itself from the standpoint of those or “the one” who has transcended such a level of intelligence? Gamez’s point of view, somewhat unusual but extremely cleverly set, is “that only a machine could think, and indeed only very special kinds of machines, namely brains and machines that had the same causal powers as brains. And that is the main reason strong AI has had little to tell us about thinking, since it has nothing to tell us about machines. By its own definition, it is about programs, and programs are not machines. […] No one would suppose that we could produce milk and sugar by running a computer simulation of the formal sequences in lactation and photosynthesis, but where the mind is concerned many people are willing to believe in such a miracle because of a deep and abiding dualism: the mind

---

27 D. Gamez, Human and Machine Consciousness, unpaginated.
28 Ibid., unpaginated.
they suppose is a matter of formal processes and is independent of quite specific material causes in the way that milk and sugar are not.”

Philosopher José Luis Bermúdez is not a pessimist when it comes to the future tasks of philosophy, from interpreting the possibilities of using the so-called non-conceptual languages to the intelligence of machines and computers. The fact that we currently know very little and not enough about the phenomenon of consciousness does not mean for him at all that the development of science will not reach such levels of problem solving in which philosophy will play a considerable role, and interpretation will not be reduced to theoretical physics, natural sciences, and technological frameworks. In his opinion, science and philosophy should work together to elucidate the nature of consciousness in all its forms, and as for artificial intelligence, that is, the development of understanding and consciousness in something outside of man, he is also optimistic.

He believes that consciousness will be reproducible in machines, computers, and robots because “consciousness already exists in a machine – the human body – so why not in an artificially created body? Of course, these have to be organic machines, and we do not know how to make organic machines for now. But I have never been inclined to easy leaps, from not having a clue how to do something at the moment to that being impossible to do.”

Another thinker sending a powerful message is Aristea Papadimitriou, who puts thoughts of the future consciousness beyond the human brain and mind, and the development of AI into the context of drastic changes that will encompass all communication aspects. Contemporary and now available philosophical and scientific outlines and predictions about the “new” intelligence will have to include a discussion of the philosophical aspects of communication and the interrelations that, until now, need not or could not have been considered. As this philosopher and communicologist has pointed out: “Within the last decade AI technology exploded and only in the last two years many advances in machine learning were made rapidly. Among the major developers and investors in AI research which also use this technology are the social networks. As AI is an interdisciplinary field it calls for the collaboration of many disciplines for the creation of its theoretical framework and its ethical code so as to stand as beneficial force for the society. As digitalization is an integral part of our culture which has already brought many alterations in our human condition, the introduction of intelligent machines to our life as active members of our society is about to seal the end of an era. The most challenging part of AI technology is that it raises again the fundamental questions about what it means to be human and, as Aristotle would suggest, reminds us not to lose sight of what a human being is as a whole.”

30 D. Gamez, Human and Machine Consciousness, unpaginated.
31 https://www.index.hr/vijesti/clanak/svjetski-filozof-za-index-slobodno-me-citirajte-roboti-ce-imati-svijest/2025515.aspx, (last accessed on December 8, 2019).
Questions about the meaning of human existence have always been linked to a search for the foundation that is at the root of everything animate and inanimate. This foundation is seen as something or someone that – standing outside human time and space – is immortal. Experience has shown that all that is biological is subject to entropy. Cybernetics as the science of the general laws guiding the processes of management and communication, and the ways of receiving and exploiting information in organized systems – machines, living organisms, and their interrelations – has offered a concept of deliverance from the danger of decay, not only of body and soul, but of all biological and social wealth. As a new metaphysics, cybernetics has offered an answer to the cause and essence of everything that exists.

In order to preserve life in its new foundation and form, biology must be improved and upgraded by scientific and technical means. Technical processes today integrate organic and inorganic nature, and man is transformed and restructured according to the needs of the cybernetic, scientific-technical mind. There are cybernetic beings, cyborgs, robots, and avatars among us who are able to perform physical and intellectual tasks that we once believed were reserved exclusively for humans. Attempts at reaching eternity as the immortality of body and mind in the apparent reality or in the mechanical body have the character of the super-human or non-human and lead towards the cybernetic, post-human scientific-technical society.

Artificial intelligence on the level of super-intelligence will create the preconditions for life beyond biology, which will have its own scientific-technical logic of development. Man’s search for the immortal and eternal, and man’s desire to come close to it, will end in the creation of a non-human body and mind, outside and beyond the space-time of a possible human history.
Bibliography:


Filozofija i svjesnost u budućnosti – kiborzi i umjetna inteligencija u iščekivanju besmrtnosti

**Sažetak**

Prirodne znanosti i tehnologije stavlja u središte ljudske pozornosti umjetnu inteligenciju, robotiku i kiborge. Međutim, virtualna i proširena stvarnost te nezamislive mogućnosti budućih medija i komunikacija između pojedinaca i društvenih skupina mogle bi biti dublje i šire nego što mislimo i razvijati se u oblicima kojima se još ne nadamo. Ako se filozofija želi orijentirati u biti znanstveno – tehničkog doba, ona mora misliti same temelje računalno – operativnog mišljenja jer umjetna inteligencija već danas razvija vlastitu svjesnost a u budućnosti i strategiju razvoja mimo čovjeka i bez čovjeka.

**Ključne riječi:** budućnost, filozofija, kiborzi, komunikacija, mediji, svjesnost, umjetna inteligencija.