ESZTERHÁZY KÁROLY KATOLIKUS EGYETEM
INFORMATIKA KAR • DIGITÁLIS TECHNOLÓGIA INTÉZET
AGRIA MÉDIA KONFERENCIA 2023

https://doi.org/10.17048/AM.2023.259
https://videotorium.hu/hu/recordings/51407

Dr. Péter Antal: Second machine age: artificial intelligence in the 21st century

Dr. Péter Antal
Eszterházy Károly Catholic University, Department of Digital Culture
antal.peter@uni-eszterhazy.hu

Abstract: While artificial intelligence has become an integral part of our life, we are still not fully familiar with the respective ethical, legal and moral issues along with its current and prospective social impact. Although the progress appears uninterrupted, certain questions related to the respective developmental levels (General AI vs. Strong AI) should still be clarified. Most researchers maintain a skeptical attitude towards the emergence and actual time of the onset of technological singularity despite the importance of being prepared when it takes place.

Artificial General Intelligence (AGI) cannot be considered either inherently good or bad. However, regarding the speed, complexity and intelligence level it will be a hardly recognizable moment when Artificial General Intelligence will surpass the capabilities of humans. Therefore, it is increasingly important that attention be paid to its reliability and to the reconciliation or synchronisation of the given objectives before its large scale introduction.

Such developments compel us to search for answers to the following questions:

How can we establish or elaborate more durable and robust AI systems than those of the current ones and which follow our commands without collapsing or being exposed to operational error or hacker attacks?

How can we make our legislative mechanisms up to date in order to promote fairness and efficiency while keeping pace with the fast changing "digital landscape?"

How can automation contribute to social well-being without depriving people of their income and a perspective for a productive life?

My lecture exploring AI's potential impact on the future of humanity searches for answers to these and similar questions.

Keywords: Artificial General Intelligence, AGI, AI, Cultural Evolution, future research

1. Introduction

While artificial intelligence has become an integral part of our life, we are still not fully familiar with the respective ethical, legal and moral issues along with its current and prospective social impact. Although the progress appears uninterrupted, certain questions related to the respective developmental levels (General AI vs. Strong AI) should still be clarified. Most researchers maintain a skeptical attitude towards the emergence and actual time of the onset of technological singularity despite the importance of being prepared when it takes place.

Artificial General Intelligence (AGI) cannot be considered either inherently good or bad. However, regarding the speed, complexity and intelligence level it will be a hardly recognizable moment when Artificial General Intelligence will surpass the capabilities of humans. Therefore, it is increasingly important that attention be paid to its reliability and to the reconciliation or synchronization of the given objectives before its large -scale introduction. Consequently, it is crucial to retrace the process during which humanity became capable of producing artificial intelligence while exploring its potential impact on our future.

2. Intelligence and the stages of life

According to Max Tegmark, the concept of intelligence can be simply defined as *the ability* to achieve *complex goals or* realize *complex objectives* [10]. The emergence of intelligence, however, was a very long process as about 4 billion years ago, at the time of the first signs of human life, intelligence in the present sense had not yet existed. We can also argue that similarly to the formation of life intelligence itself had undergone an unique evolutionary process as well.

In a broader sense life can be considered a *self-multiplying information system capable of maintaining its own complexity*. [10] Accordingly the growth of life can be divided into three stages.

Life 1.0, biological evolution (biological)

Life 2.0, cultural evolution (humans - learning)

Life 3.0, technological evolution (beyond biological intelligence)

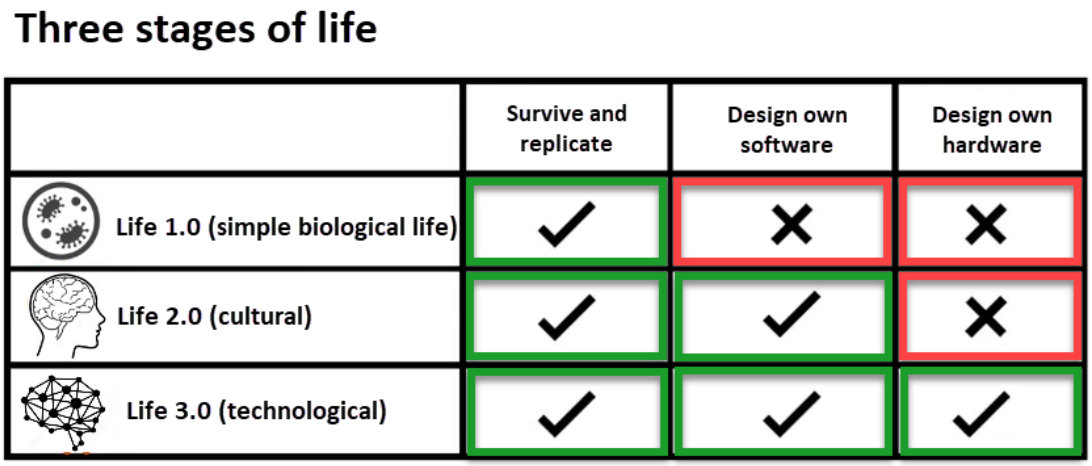

Figure 1 *The main features of the three stages of life (author's own figure)*

Life 1.0 is the biological evolution during which the first organisms copied gene-determined information required for multiplication. As a result of this process it wasn't the matter composed of atoms that multiplied, but information containing bits and determining the arrangement of atoms. Accordingly, if a bacterium copied its own DNA, new atoms were not formed, but an additional line of atoms was arranged into a pattern identical to the original.

In response to environmental influences mutations emerged in this genetical code eventually driving the development process. Such growth, however, was extremely slow as both the hardware and software result from evolution instead of design, in other words, behaviour (software) and body structure (hardware) were determined by information stored in the DNA.

Life 2.0, or cultural evolution, is best exemplified by the human being, and the main difference is that the respective software is a result of design instead of evolution.[6] One tremendous advantage is that the human hardware (the body) goes through a developmental process while software (knowledge) is the result of learning, which eventually means that the intelligence obtained in this manner is not dependent on the quantity of information stored in the DNA.

[6] Software includes the sum of algorithms and knowledge we use in processing information obtained via the sensory organs and rely on the subsequent decision-making process. The term entails such abilities as the recognition of familiar faces, motion, reading, writing, counting, singing, and telling jokes.

The synaptic connections linking our neurons can store approximately hundred thousand times more information as the DNA at the time of birth. Our synapses store all knowledge and skills in about 100 terabytes, while the DNA can only carry 1 gigabyte of information. Life 2.0 is more flexible and smarter than Life 1.0. The latter can adapt only slowly to the changes in the environment via the gradual evolutionary process encompassing generations. Bacteria that encounter antibiotics can become resistant to certain medicines only after several generations. Life 2.0, however, is capable of fast adaptation to the given environment via the refreshing or updating of the respective software. One such example is that if someone recognizes sensitivity to lactose, they can change their behavior and will avoid consuming dairy products.

Right after birth, humans can only perform reflex-type activities, and during their growth, they will learn new, complex skills. At first parents, later the school, then the humans themselves determine what type of ”software” will be encoded into their hardware. While in the past 50,000 years the quantity of information stored in the human DNA had not changed significantly, the amount of information or knowledge recorded in the human brain, books, and computers had grown by magnitudes.

Such explosive development is due to the flexibility of the brain facilitating the acquision of skills leading to the formation of intelligence. This built-in software module enabled the installation of such ”tools” as the capability of multichannel communication enabling the copying of information stored in the brain of another person thereby allowing information to survive after the death of the brain. The software module of reading and writing allowed saving information into external storing systems enlarging the original brain capacity via storing and retrieving knowledge. As Marshall McLuhan asserts, knowledge today does not mean the possession of information, but the ability to navigate or orient oneself in the incessant flow of information transmitted by electronic means [7].

Although humanity has started on the road toward the implementation of the Life 3.0 stage and certain updating of hardware is possible via artificial body parts such as limbs or the hearts neither the human lifespan can be extended to hundreds of years, nor the size of the brain can be increased.

The **Life 3.0** stage advances even further, allowing the system to gain total mastery. By gaining full control not only over its software but hardware too, it would be free from the limits inherent in evolution, eventually ushering in the *age of Artificial General Intelligence* (AGI,) whose most important features are the following:

it exceeds the stage of biological intelligence,

it can respond to environmental influences without continuous human interference (automatic devices)

it can behave similarly to an organism possessing natural intelligence (machine-based characters)

it can change its behavior deliberately and in a repeated manner (learning, self-development).

3. The phases of cultural evolution and its future

The history of the evolution of the human mind and the education process can be considered a continuum representing the passing and processing of information via interpersonal communication between several generations [2]. At the same time the structure of knowledge had undergone multiple changes driven by new forms of communication and the advance of technology. This process is the cultural evolution, which exceeds the biological evolution stage of the development of the human community. The long way leading to this point retraced below was not without obstacles and challenges.

Episodic culture

The basic difference between the human and animal mind is the development of the brain. In case of animals the responses to environmental stimuli are recorded in the genetic code, which changes during the evolutionary process while the best responses given to environmental changes are selected and stored in this location.

In the case of some animals and humans as well, experience acquisition on an individual basis apart from the genetic code became possible, leading to the formation of abstractions regarding the given environment. All this

resulted in the development of the brain, which as an open system, becomes programmed by environmental influences via the accumulation of experiences.[7]

Culture transfer

As one of the requirements of cultural evolution, special forms of information transmission, known as culture transfer, had to be performed. Compared to genetic transfer, this process facilitates the gradual taking over or adoption of the experiences and ”knowledge” of other members of the same species. In the case of humans, the forms of culture transfer include mimetic learning, knowledge acquisition via instruction, and learning through cooperation.

Sociogenesis, implying culture transfer among generations, is a unique cognitively constructed human feature.

Mimetic culture

The first paradigm shift during human evolution was the transition between episodic and mimetic culture. As Merlin Donald argues, preceding the appearance of human languages, mimetic culture emerged approximately 1,5 million years ago with the Homo erectus [3]. Mimetic culture can be defined as a deliberate or intentional representation by means of one’s body, characterized by intentionality, openness, and the awareness of communicational intent. The starting point was the self-projection of one’s body, or, in other words, imitation. The emergence of mimetic culture led to several developments, including the ”distribution” of concepts, the prevalence of the social game system, the appearance of organized, formal means of ”education,” coordination via group communication and shared representation, or the ability to renew coordination as manifested in such activities as hunting [9].

Cooperation with peers, shared thinking, and social attraction generates communicational compulsion. The concept elaborated by Csányi [2] implies the human specie-specific means of knowledge transfer and reception via instruction and learning.

The evolutionary design of pedagogy suggests that humans even at the time of birth ”are aware” of the importance of adults in their immediate vicinity as valuable source of knowledge. Such forms of behavior function as components of our genetic heritage and are integrated into our cognitive architecture [3].

Mythic culture

One form of successful adaptation potentially utilizing certain prototypes of language use is mimetic culture. Accordingly, complex forms of social activity developed, mostly due to the shared foundation of an episodic database closely connected to events and situations could only change slowly [3).

The emergence of speech facilitated more effective forms of communication, and a new level of model construction and symbolic invention (ingenuity, or resourcefulness) appeared. The representation of reality with linguistic symbols provides an extremely effective modeling and communication device while facilitating exact and ”broadbandwidth” information transmission, functioning as the ideal medium of instruction and learning. During the speech process the acoustic signs are complemented with visually perceptible metacommunicational, non-verbal signals. Furthermore, the emotional aspect related to communication can be ”broadcast or conveyed” via acoustic signs as well.

Theoretic culture

Theoretic culture emerges simultaneously with the discovery of external storage devices. The first such development was the appearance of writing, which led to the formation of an external storage system. In previous cultures, knowledge was held in the mind of a given individual, and via writing, it became available to anyone. Donald named this publicly available or accessible system, External Symbolic Storage (ESS) [3]. The rise of ESS

[7] It is important to note that at this stage of cultural evolution learned abstractions could not be transferred to the next generation.

made human knowledge transmission independent of person, space, and time. Information transmission did not require the presence of the sender and receiver, leading to the new dimension of transmissible and receivable knowledge. Literacy resulted in changes of memory as functioning in the new system requires only the knowledge of the exact location and form of new information. The rise of ”theoretic” culture facilitated the appearance of a higher level and more effective analytical thinking and theory formation.

Digital culture (the age of electronic media and networks)

The development of technology changed the information environment in two distinguishable ways. One tendency is the ”placement” of the operational capability of the brain into external memory systems used for the storing and re-processing of information.

The other determinative factor is the development of telecommunication and information technology. The process driven by such dramatic changes as the emergence of media convergence and the rise of a new network infrastructure resulted in the prevalence of digitally provided internet services. This new information technology constellation gave rise to a new information-based reality potentially changing our cognitive architecture, the arrangement of our representations and our cognitive habits as well.

Life 3.0

The onset of the information explosion can signifiy the arrival of technological singularity. Technological singularity is a hypothetical event which occurs when the strong artificial intelligence (AGI) reaches a point of development when human intelligence and technology is exceeded by a magnitude incomprehensible by previous (pre-singularity) thinking.

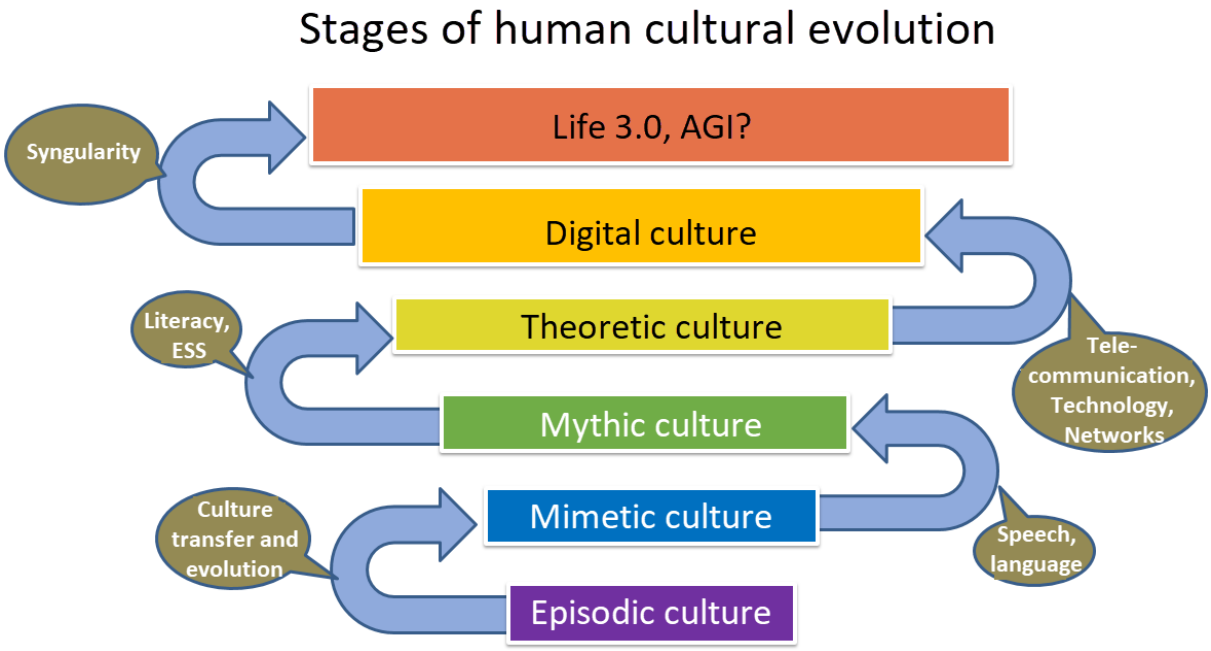

Figure 2 *The stages of human cultural evolution (author's own figure)*

4. The debate concerning the future of artificial intelligence

Humans in today's sense have appeared on Earth hardly more than 100 000 years ago, thus as far as the evolutionary process on Earth is concerned, Life 2.0 has been around only for a relatively short time. The rise of Life 3.0, that is, of the artificial general intelligence, (AGI)[8] is a rather controversial subject. Several researchers predict

[8] AGI is such an intellingence, which is capable of realizing any goal or objective, including learning.

that AI will surpass human intelligence in this or the following century. The way of the actual onset and the respective impact are of general concern.

The two most important questions are ”the when and the how.”

When will the artificial general intelligence (AGI) reach and surpass the level of human intelligence, in other words when will Life 3.0 emerge? Will this take place in our life or only a few hundred years later?

What will its impact be on humanity? (will this process be beneficial for us?) and how will it take place? (What kind of social, economic, legal, or military consequences can we expect?)

Researchers and experts haved identified three clearly separable approaches regarding the abovementioned questions:

Digital utopians,

Technosceptics,

Beneficial AI movement.

Both the technosceptics and the followers of the digital utopia perspective tend to diminish the danger based on different reasons. The technosceptics believe that no artificial intelligence can emerge in the foreseeable future, while digital utopians are convinced that the respective point of development will be achieved, but it will only have advantages. The proponents of the Beneficial Intelligence perspective share the given concern based upon relevant research and scholarly discussions. The new luddites[9], the destroyers of machines are naturally afraid of such developments and reject any form of AI.

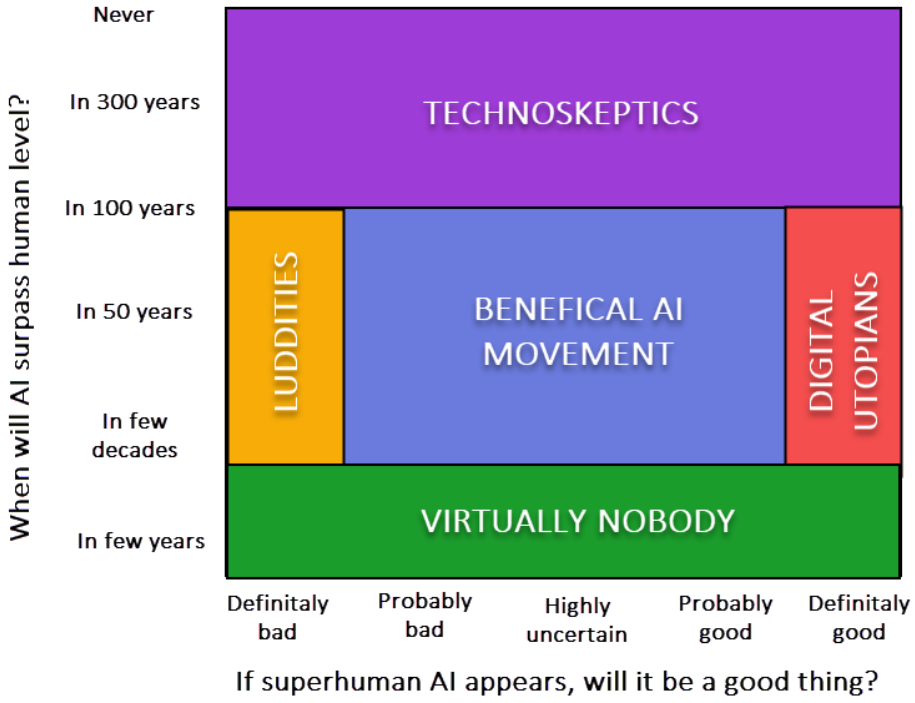

Figure 3 *Philosophical tendencies concerning the future of AGI [*10*]*

9 The term originally refers to followers of Ned Ludd, a legendary figure resisting the Industrial Revolution in England in the late 18th and early 19th century. The main forms of resistance was destroying machines.

While the truth must be in-between, it is encouraging that discussions started concerning the most important questions. The most reasonable group is the Beneficial AI movement. One of its spokesmen is the astrophysicist, Max Tegmark, who is also the leader of the Future of Life Institute.[10] The organization prioritises preparation for the potentially negative consequences in order to enable humanity to prevent the problems while maximising its advantages and avoiding the related pitfalls. The Institute also aims to preserve the favorable impact of AI on society, to motivate research in several fields ranging from the economy and law to such technology-dominated aspects as monitoring, validation, security, and control.

5. Doubts and facts

As a result of the learning process human intelligence operates on a broad spectrum, as we can acquire any kind of knowledge, while our brain can only operate within set physiological limits. The hypothetical concept of information explosion is based on the belief of the emergence of AI matching the human level. This new form of intelligence is free of all biological limitations and is capable of self-development. Due to the fast feedback computations of any magnitude within a given time period would be possible as well. One of the greatest deficiencies is that current AIs are not universal. While they are capable of learning, the knowledge acquisition process is reduced to certain areas and is within certain limits and follows pre-set trends. The connection between universality and computing capacity is highly debated.

Irwing J. Good had already declared in 1965 that ”since ultra-intelligent machines are expected to surpass human intelligence, such machines would be the last invention humans should come up with” [4].

Raymond Kurzweil the noted AI researcher and the author of The Age of Spiritual Machines: When Computers Exceed Human Intelligence, and The Singularity is Near: When Humans Transcend Biology predicts the onset of singularity for 2045. While the expected date appears to be close, it is due to the linear illusion of development, since the actual process is believed to take place in an exponential manner. Accordingly, if measured in present terms, the 21st century will see not 100 year growth, but a developmental pace of 20 000 years. [5]

Kurzweil justifies his prediction by the principle of exponential development seen in several technological fields. One of the best known theories is the Moore Law [8] according to which the complexity of integrated electric circuits doubles on an 18 month basis. Similar exponential growth can be expected in case of the speed of computation or the size of chips.

6. Options and challenges

The development of AI is a crucial and strategically important concern virtually in all economic sectors including transport, energy industry, communication and media, military industry, health care, lawmaking or the processing of languages.

Nevertheless, major issues have to be clarified before its widespread use as such programs can impact the basic aspects of human existence. The argument that AI has no objective, intuition, creativity, or language steadily loses validity. Consequently, reliable and accurate answers have to be given to the following questions in the near future:

How can we establish or elaborate more durable and robust AI systems than those of the current ones and which follow our commands without collapsing or being exposed to operational error or hacker attacks?

How can we produce smart weapons while avoiding casualties of innocent civilians and precluding the possibility of an uncontrollable arms race in the field of automatic weapons?

How can automation contribute to social well-being without depriving people of their income and a perspective for a productive life?

Security is the most important concern. Humanity has always tried to learn from its mistakes throughout their development as heretofore achieved technological milestones had more advantages than disadvantages. In order to retain control over the AGI in the future researchers of AI security call attention to four aspects of AI software:

[10] Future of Life Institute: https://futureoflife.org

Verification – the elimination of BSOD (blue screen of death) (was the system appropriately constructed?)

Validation – robots should learn to be careful, (have we built the right system?)

IT security – continuous and full protection of the integrity and availability of data and devices,

Control – the autonomy of machines, controllability.[11]

7. Social challenges

As far as society is concerned there are several serious issues related to the impact of AI. How will AI influence our work, the labour market, the incomes, what kind of trade or profession should the new generations be trained for?

Erik Brynjolfsson, the father of the Digital Athens theory positively views the connection between humans and the AI. He believes that just like in ancient Athens where most labor was performed by slaves enabling the people to spend more time with arts, sports, etc, machines can be enslaved and serve as the foundation for the age of abundance, stress-free life, and consumption [1]. Although such a social and economic arrangement appeared to work in the USA until the 1970s, the distribution of income was not equal. While the economic growth continued in the following decades, incomes started to stagnate for 90% of the population and the respective surplus became owned by the superrich 1%.

Although economists find the reasons for such dramatic changes in the global market trends, certain researchers blame technological development. Accordingly the growth of technology contributed to the inequality in three ways.

Old established professions became obsolete and the new occupations requiring more skills and expertise became available only to the highly qualified segments of society.

Due to automatization most income is realized at the owner or enterprise level. Accordingly, in 1990, the Big Three of Detroit (Chrysler, Ford, GM) realized an income identical to that of the Big Three of Silicon Valley (Google, Apple, Facebook), yet the latter employs nine times fewer people and has a thirty times greater value at the stock market [6].

The digital economy is star -oriented, as one application can become predominant over the others.

Yet, one thing is sure, learning is profitable and in the future better paid positions will require human interaction, social intelligence, creativity, and fast problem solving skills along with the ability to adapt to the varying and unpredictable work environment.

8. Present and future perspectives

AI raises several questions, and it is up to us whether our future will be suitable for humanity or it will prove to be fateful. Currently, the most important task is the setting and synchronizing of objectives. This process includes the following components:

A widespread and comprehensive agreement has to be reached concerning the acceptance of AGI objectives (freedom, use) and their application in case of the rights of non-humans (animals).

The synchronization of AGI with human objectives implies three challenges:

AGI should be aware of human objectives in order to **understand t**he underlying reasons, hypotheses, sub-objectives, and priorities.

The acceptance of general human goals and priorities, i.e. the value system behind raising children.

[11] source: AI Safety Resarch https://futureoflife.org/landscape/

A guarantee that AGI will **honor and preserve** human goals even though the former will surpass the latter in knowledge and skills.

The future is uncertain, and it is possible we do not reach the stage of 3.0, or technological singularity will mark the end of humanity. The potential outcomes are summarised by the following info-graph.

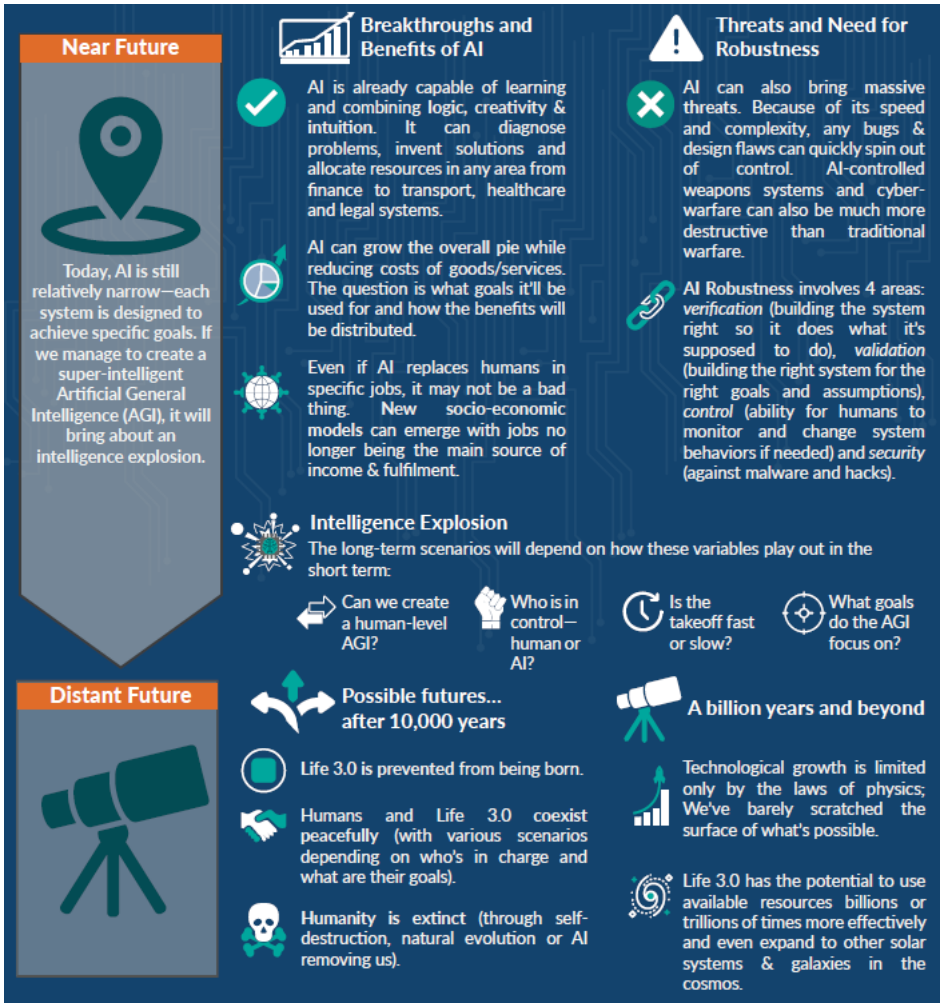

Figure 4 *Potential* Outcomes *of Life 3.0*[12]

9. Summary

It is beyond doubt that AI can surpass human intelligence, but its behavior cannot be predicted. Humanity has not yet become prepared to avoid the potential pitfalls, the intentional or unintentional tricks the system can play on us. One of the best examples could be found in evolution. Humans are in control of the planet not because of being the strongest, fastest, or the largest, but because of being the smartest. Thus the question emerges, what will happen when we humans will not be the smartest on the planet, can we still retain control over the machines?

[12] Source: https://readingraphics.com/product/download-life-3-0-summary-in-pdf-audio-graphic/ Copyright 2018. Skool of Happiness Pte Ltd.

WORKS CONSULTED

Brynjolfsson E., Mcafee A. (2014). The Second Machine Age: Work, Progress and Prosperity in a Time of Brilliant Technologies, New York, Norton, 2014. ISBN13: 9780393350647

Csányi, V. (2006). Az emberi viselkedés, Samona, Nők Lapja Könyvműhely, Budapest, 2006. ISBN 9639710075

Donald M. (2001). Az emberi gondolkodás eredete. (Origins of the Modern Mind: Three Stages in the Evolution of Culture and Cognition). Osiris Kiadó, Budapest, 2001. ISBN: 963-389-085-3

Good, I. J. (1965). „Speculations Concerning the First Ultraintelligent Machine", in: *Advances in Computers* vol 6, Franz L. Alt and Morris Rubinoff, eds, 31-88, 1965, Academic Press

Kurzweil, R. (2013). A szingularitás küszöbén. Amikor az emberiség meghaladja a biológiát. (The Singularity is Near: When Humans Transcend Biology) Ad Astra kiadó 2013. ISBN 978-615-5229-26-8

Manyika, J. „Can technology and Productivity save the day?" presentation. https://futureoflife.org/data/PDF/james_manyika.pdf?x17135 download time: 21-03-02.

McLuhan, M. (2011). The Gutenberg Galaxy. Toronto: University of Toronto Press

Moore G. E. (1965). Cramming more components onto integrated circuits In: *Electronics*, Volume 38, Number 8, April 19, 1965. https://newsroom.intel.com/wp-content/uploads/sites/11/2018/05/moores-law-electronics.pdf download time: 21-03-15.

Pléh, CS. (2001). Tudástípusok és és bölcsésztudományok helyzete, In: *Világosság* 2001. (42. évf.) 7-9 sz. 11-30.

Tegmark, M. (2018). Élet 3.0 Embernek lenni a mesterséges intelligencia korában (Life 3.0: Being Human in the Age of Artificial Intelligence) HVG-Könyvek, Budapest, 2018. ISBN 978-963-304-504-6