

Freedom and Truth in the Age of Artificial Intelligence

Milan M. Milosavljević

Abstract— In this paper, we address the Age of Artificial Intelligence, which we define as a distant future in which humans will reach the HLAI (Human Level of Artificial Intelligence) point. How humans will relate to fundamental concepts, such as freedom and truth. What will happen with today's basic human drivers, such as the death drive, anxiety and enthusiasm? Will they remain as basic ways of our entanglements with Being? We are convinced that the answering these questions today have full meaning not only for understanding the future, but primarily because of our understanding of modern human society and our openness toward new possibilities.

Index Terms—HLAI, Truth, Freedom, Entanglements with Being.

I. INTRODUCTION

After sixty years of development, the field of Artificial Intelligence (AI) is dominated by two very different goals. The first one is the development of intelligent systems, which may or may not be directly related to a human intelligence, his way of thinking and behavior. Second one, ultimate radical goal, is to reach the level of human intelligence within the framework of various types of artificial systems. This approach is newly reformulated as a separate branch of AI – so called HLAI (Human Level of Artificial Intelligence), See Fig.1, [2], [7]. In section II we will analyze some sufficient conditions for HLAI.

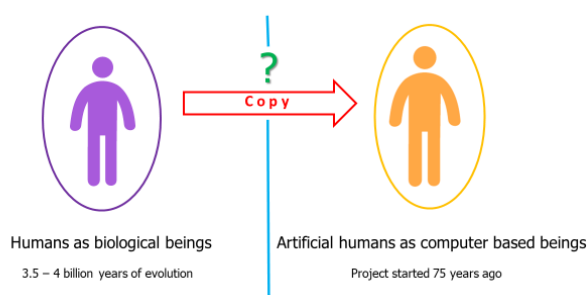


Fig. 1. Fundamental problem of Human Level Artificial Intelligence.

We will point to an important Kant's observation regarding the moral attitude of people to whom would be available noumenal world. It turns out that this attitude is very applicable in the circumstances of the asymptotic reaching HLAI. In section III we will consider the consequences of producing intelligent system without

Milan Milosavljević is with the Singidunum University, Danijelova 32, 11000 Belgrade, Serbia (e-mail: mmilosavljevic@ singidunum.ac.rs), and also with the Faculty of Electrical Engineering, Belgrade University, Kralja Aleksandra 73, 11000 Belgrade, Serbia

conscious. It turns out that these systems cause a fundamental shift of human authority to authority of data and algorithms. This is beginning of new age in human history, which is, by some authors, called Dataism. In concluding section we briefly summarize open dilemma and opportunity of 21. century AI.

II. HUMAN LEVEL OF ARTIFICIAL INTELLIGENCE

Sufficient condition for the achievement of this objective is the copying of humans from their natural biological substrata to some chosen computing base. The copy operation implicitly includes several conditions that at first glance remain hidden to us.

- Homo Sapiens can be accurately modeled by a computable class of mathematical mappings.
- Information capacity of copies is not less than the information capacity of Homo sapiens.
- Information content of the sensor inputs of a copy is not smaller than the information content of the sensor inputs of Homo sapiens.
- Internal states of a copy are identifiable in relation to the complete internal states of Homo sapiens.
- The evolutionary potential of a copy is not less than the evolutionary potential of Homo sapiens.
- Intentionality of mental processes of a copy are identical to the intentionality matrix of Homo sapiens.
- Social, historical and spiritual horizon of a copy coincides with the social, historical and spiritual horizon of Homo sapiens.
- Language structure of a copy is equal to the linguistic structure of Homo sapiens language.
- Exposure of a copy to the natural, historical, cultural and sociological factors of the environment is not less than the same exposure of Homo sapiens.

The above statements are only one illustrative set of preconditions for successful copy process. It is difficult to say whether they are the minimal or necessary.

But already in this illustrative form it gives us clear idea of the complexity of the copy process and potential differences between the copy and the original, although the original and the copy was verified as identical within some chosen computable formal system.

Note that the corps of knowledge and technology, necessary for the process of copying Homo sapiens, at the same time would be sufficient for the creation of a whole class of artificial Homo sapiens, whose characteristics in certain segments significantly exceeded the capabilities of

the original. For example, by increasing sensory or cognitive capacity or by embedding previous experience inaccessible to every single original.

Let label man-made Homo sapiens generation as a 0-th generation, which is different from the original Homo sapiens at least in some important properties (cognitive, sensory, motor, emotional, social...). The emergence of 0-th generation would lead to rapid techno - evolutionary progress of each subsequent generation.

The question of coexistence of generations labeled by different ordinal numbers, would be equivalent to the question of the practical usefulness of the steam-powered cars compared to the modern gasoline or electric drive cars. Each succeeding generation would making the previous generation radically unnecessary.

Is the exponential rise of the new Homo sapiens really guaranteed?

History teaches us that the natural Homo sapiens is result of complex, especially social relations, and harmonious integration into the current ecosystem. In recent 70 000 years, a dramatic rise of Homo Sapiens is not the result of any improvement of its biological substrate, but rather the result of development of cognitive abilities arising from the complexity of human communities and the relationships within it. The dominant evolutionary forces are no longer dictated by the environment, but also by social, economic and communications structure of the human community. The greatest enemy of the Homo sapiens are no longer the forces of nature and predators, but other Homo sapiens.



Fig.1 Immanuel Kant (1724 – 1804) was a German philosopher who is considered a central figure in modern philosophy. In a mysterious subchapter of his *Critique of Practical Reason* titled “Of the Wise Adaptation of Man’s Cognitive Faculties of His Practical Vocation,” he try to answer the question of what would happen to us if we were to gain access to the noumenal domain.

It is logical to assume that the dominant forces of survival of each new generation of Homo sapiens will have a source in their own community. Additional complexity and incomparability with the present time is the existence of essential centers of power who make decisions about the birth of new generations. In this new world of Homo

sapiens, individualism, free will and the connections with other community members will have crucial impact on their survival.

In projecting the futuristic vision of Homo sapiens it is hard to avoid our today's position and the restrictions imposed by the actual socio-economic and historical moment.

Will the new Homo sapiens follow our today’s mental matrix?

In this respect, it is inspiring to consider Kant's analysis of hypothetical situation in which humanity would be if it will have access to noumenal world, Fig.1. For the type of consideration carried out in this work, Kant's noumenal world is similar to the world at the time of appearance of 0-th generation of artificial Homo sapiens. We find this interesting passage in a mysterious subchapter of his *Critique of Practical Reason* titled “Of the Wise Adaptation of Man’s Cognitive Faculties of His Practical Vocation,” in which he try to answer the question of what would happen to us if we were to gain access to the noumenal domain, to the thing in itself (Ding an sich):

“Instead of the conflict which now the moral disposition has to wage with inclinations and in which, after some defeats, moral strength of mind may be gradually won, God and eternity in their awful majesty would stand unceasingly before our eyes.... Thus most actions conforming to the law would be done from fear, few would be done from hope, none from duty. The moral worth of actions, on which alone the worth of the person and even of the world depends in the eyes of supreme wisdom, would not exist at all. The conduct of man, so long as his nature remained as it is now, would be changed into mere mechanism, where, as in a puppet show, everything would gesticulate well but no life would be found in the figures.” [4].

In the Žižek work *Organs without Bodies: On Deleuze and consequences*, we found interesting interpretation of this passage:

“So, for Kant, the direct access to the noumenal domain would deprive us of the very spontaneity that forms the kernel of transcendental freedom: it would turn us into lifeless automata, or, to put it in today’s terms, into “thinking machines.” [5].

We illustrate this strange situation in Fig.2. Approaching the age in which the HLAI is possible, we are simultaneously approaching the era when we gain access to the noumenal domain. On this path we humans are inevitably transformed into creatures like thinking machine, so that the appearance of 0-th generation of new Homo sapiens can be seen as an asymptotic generation of natural Homo sapiens.

At the level of phenomena but also at the noumenal level, humans are a “mere mechanism” with no autonomy and freedom. As phenomena, we are not free. We are a part of nature, a “mere mechanism,” fully submitted to causal chains. As noumena, we are again not free but reduced to a “mere mechanism.” Our freedom exists only in a space in

between the phenomenal and the noumenal, Fig.3.

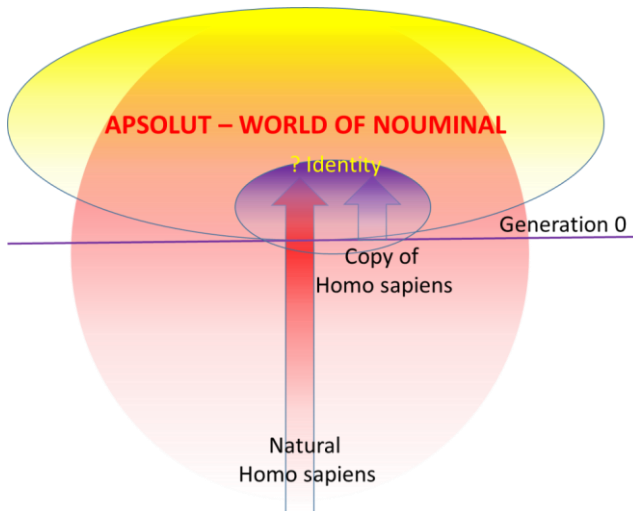


Fig.2 Approaching the age in which the HLAI is possible, we are simultaneously approaching the era when we gain access to the noumenal domain. On this path we humans are inevitably transformed into creatures like thinking machines, so that the appearance of 0-th generation of new Homo sapiens can be seen as an asymptotic generation of natural Homo sapiens.

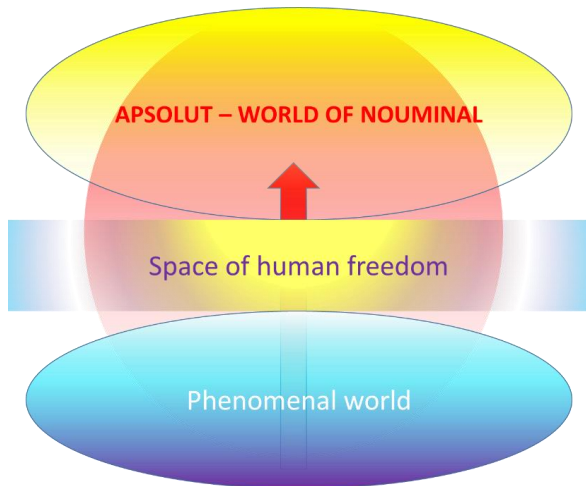


Fig.3 The place of human freedom in the gap between noumenal and phenomenal world.

III. DATAISM – INTELLIGENCE WITHOUT CONSCIENCE

When we talk about another dominant purpose of artificial intelligence, we actually talk about synthesis of intelligent systems that do not necessarily follow the idea of human ways of solving problems. Recently, a significant shift has been made along the development line of triad: Big Data - Machine Learning - Data Science. The basis of this progress is the fact that the enormous generation of data in modern global computer communications networks enables successful operation of machine learning algorithms [8]. The performance of these systems are begin to beating the evolutionary developed natural system, but in a much shorter development time (training).

All this has led to the shift of human value systems and authority.

For thousands of years humans believed that authority came from the gods. Then, during the modern era, humanism gradually shifted authority from deities to people. Jean-Jacques Rousseau summed up this revolution in novel *Emile, or On Education*. When looking for the rules of conduct in life, Rousseau found them “in the depths of my heart, traced by nature in characters which nothing can efface. I need only consult myself with regard to what I wish to do; what I feel to be good is good, what I feel to be bad is bad” [6]. Humanist thinkers such as Rousseau convinced us that our own feelings and desires were the ultimate source of meaning, and that our free will was, therefore, the highest authority of all [3].

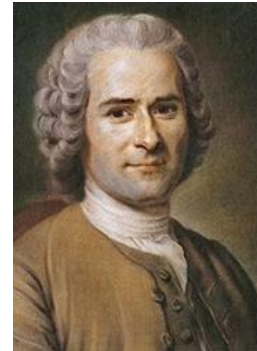


Fig. 4 Jean-Jacques Rousseau (1712–1778) Francophone Genevan philosopher, writer, and composer of the 18th century. His political philosophy influenced the Enlightenment and the overall development of modern political and educational thought. In novel *Emile, or On Education* he wrote: “I need only consult myself with regard to what I wish to do; what I feel to be good is good, what I feel to be bad is bad.”

The individual is the source of all meaning and authority. The individual is autonomous and possesses inner unity. The supreme value is liberalism. It is based on our Individual inner voice and our inner freedom. Only we can really know ourselves. We know ourselves better than any external system can know us. This principle can be extended to all fields of human social life. Liberal politics means that the voter knows best. Liberal economics demand that the customer is always right. Liberal aesthetics forces attitude that beauty is in the eye of the beholder. Liberal ethics principle is what I feel to be good, is good. Eventually, liberal education sends to us the most important message: think for yourself [3].

But liberalism will not hold any more. Google and Facebook know us better than we know ourselves [9]. Authority will shift from humans to computer algorithms. In its extreme form, proponents of the so called Dataism worldview perceive the entire universe as a flow of data, see organisms as little more than biochemical algorithms and believe that humanity’s cosmic vocation is to create an all-encompassing data-processing system — and then merge into it [3].

To make an intelligent external system to know us better than we know ourselves, its information input must be very rich. In addition to various biometric sensors monitoring our internal state, we need a systematic access to our so-called digital footprint - all the information that is left when we interacting with the global computer network. If

unconscious intelligent systems start reacting and making decisions on our behalf, it will be the first time in the Homo sapiens history that closed loop: the external world - sensor inputs - human decisions - actions influencing the external world, will be broken.

This situation could be viewed as a kind of the emergence of insulating shell around the Homo sapiens, which prevents direct communication with the environment, see Fig.5 and Fig.6. It looks like the price to be paid for the unnatural separation of intelligence and consciousness, firmly connected inside the mind of Homo sapiens. The consequences of loss of direct communication with the environment (“disconnected” humans) is difficult to predict at this point. But, if the communication is fundamental building block of natural evolution, it is clear that breaking it we suspend human evolution, at least in the form known to us up to now.

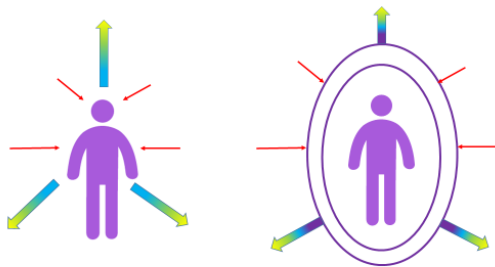


Fig.5 Influence of Internet of Things (IoT) to human exposition to outside world. It leads to so called “disconnected” humans, without closed loop flow: the external world - sensor inputs - human decisions - actions influencing the external world

What will happen with today's basic human drivers, such as the death drive, anxiety and enthusiasm? [1]. Will they remain as basic ways of our entanglements with Being? Maybe this questions will be answered by algorithms, not by “disconnected” humans.

IV. CONCLUSION

In this paper we analyze the impact of the artificial intelligence technology to the further development of Homo sapiens. In the case of HLAI as well as intelligent unconscious artificial intelligence systems, we have identified the opposing forces, which make the final outcome uncertain. There is an evident shift of authority from the humans to the data and algorithms. Descartes' *Cogito ergo sum*, would be now reads as: I exist if I generate and process data. In a similar way can be reformulated other basic concepts. The truth is not only the adequacy or revealing, it is predominantly the generation of adequate data. Freedom is no longer based on our free will, but on the adequacy of collected data and power of available algorithms which make decisions on our behalf.

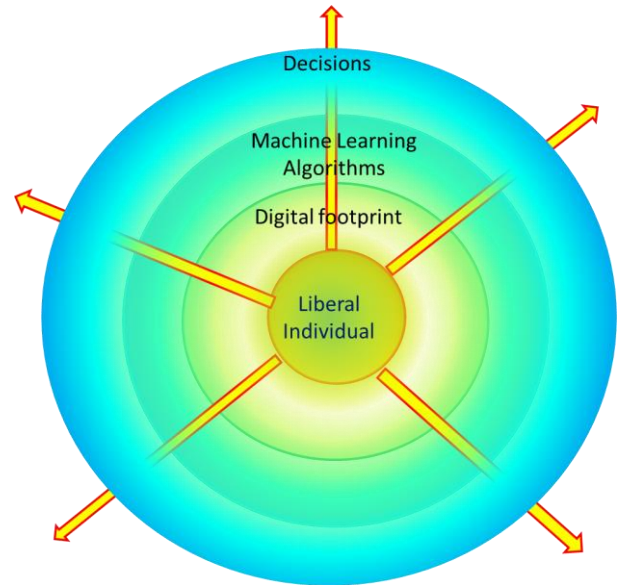


Fig.6 Structure of shells around Homo sapiens (Liberal individual) immersed in nowadays liberal capitalistic ideology, global market and artificial intelligence technology.

ACKNOWLEDGMENT

This work is partially supported by the Ministry of Education and Technological Development of the Republic of Serbia, trough project TR32054.

REFERENCES

- [1] Sigmund Freud, Civilization and Its Discontents, German title: Das Unbehagen in der Kultur ("The Uneasiness in Culture"), 1930.
- [2] Milan Milosavljevic, *Veštačka inteligencija*, Univerzitet Singidunum, 2015.
- [3] Yuval Noah Harari, Homo Deus: A Brief History of Tomorrow, 2016
- [4] Imanuel Kant, *Critique of Practical Reason*, Classic Books International, 2010
- [5] Slavoj Žižek, *Organs without Bodies: On Deleuze and consequences*, Routledge, New York and London, 2004.
- [6] Rousseau, Jean-Jacques Rousseau. *Emile, or On Education*. Trans. Allan Bloom. New York: Basic Books, 1979.
- [7] Milan Milosavljevic, Relationship between radically new and artificial intelligence, *In Proceedings of 3rd International Conference on Electrical, Electronic and Computing Engineering, IcETRAN 2016*, Zlatibor, Serbia, June 13 – 16, 2016, ISBN 978-86-7466-618-0
- [8] Milan Milosavljevic, Deep Learning and Transcendental Imagination, *Proceedings of 2nd International Conference on Electrical, Electronic and Computing Engineering IcETRAN 2015*, Silver Lake, Serbia, June 8 – 11, 2015, ISBN 978-86-80509-71-6
- [9] Wu Youyoua, Michal Kosinskib, David Stillwella, Computer-based personality judgments are more accurate than those made by humans, *PNAS*, vol. 112, no. 4, January 2015.